MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 115 No. 3 1955

GENERAL INDEX to the PUBLICATIONS of the ROYAL ASTRONOMICAL SOCIETY

A further volume of the GENERAL INDEX has now been published. It includes papers in all the series of publications of the Society up to 1950 which have not previously been included in a General Index :—

Monthly Notices	Vols. 92–110 (1932–1950)
Memoirs	Vols. 61– 66 (1917–1950)
Occasional Notes	Nos. 1– 12 (1938–1950)
Geophysical Supplement	Vols. 1– 5 (1922–1949)

The price of this volume of the General Index is One Pound; the price to Fellows is Fifteen Shillings.

Orders for copies should be addressed to :—

THE ASSISTANT SECRETARY,
Royal Astronomical Society,
Burlington House,
London, W.1.

BURLINGTON HOUSE

LONDON, W. 1

Price Thirteen Shillings and Sixpence

NOTICE TO AUTHORS

1. *Communications.*—Papers must be communicated to the Society by a Fellow. They should be accompanied by a summary at the *beginning* of the paper conveying briefly the content of the paper, and drawing attention to important new information and to the main conclusions. The summary should be intelligible in itself, without reference to the paper, to a reader with some knowledge of the subject; it should not normally exceed 200 words in length. **Authors are requested to submit MSS. in duplicate. These should be typed using double spacing and leaving a margin of not less than one inch on the left-hand side. Corrections to the MSS. should be made in the text and not in the margin.** Unless a paper reaches the Secretaries more than seven days before a Council meeting it will not normally be considered at that meeting. By Council decision, MSS. of accepted papers are retained by the Society for one year after publication; unless their return is then requested by the author, they are destroyed.

2. *Presentation.*—Authors are allowed considerable latitude, but they are requested to follow the general style and arrangement of *Monthly Notices*. References to literature should be given in the standard form, including a date, for printing either as footnotes or in a numbered list at the end of the paper. Each reference should give the **name and initials** of the author cited, irrespective of the occurrence of the name in the text (some latitude being permissible, however, in the case of an author referring to his own work). The following examples indicate the style of reference appropriate for a paper and a book, respectively :—

A. Corlin, *Zs. f. Astrophys.*, **15**, 239, 1938.
H. Jeffreys, *Theory of Probability*, 2nd edn., section 5.45, p. 258, Oxford, 1948.

3. *Notation.*—For technical astronomical terms, authors should conform closely to the recommendations of Commission 3 of the International Astronomical Union (*Trans. I.A.U.*, Vol. VI, p. 345, 1938). Council has decided to adopt the I.A.U. 3-letter abbreviations for constellations where contraction is desirable (Vol. IV, p. 221, 1932). In general matters, authors should follow the recommendations in *Symbols, Signs and Abbreviations* (London : Royal Society, 1951) except where these conflict with I.A.U. practice.

4. *Diagrams.*—**These should be designed to appear upright on the page, drawn about twice the size required in print and prepared for direct photographic reproduction except for the lettering, which should be inserted in pencil. Legends should be given in the manuscript indicating where in the text the figure should appear.** Blocks are retained by the Society for 10 years; unless the author requires them before the end of this period they are then destroyed.

5. *Tables.*—**These should be arranged so that they can be printed upright on the page.**

6. *Proofs.*—Costs of alteration exceeding 5 per cent of composition must be borne by the author. Fellows are warned that such costs have risen sharply in recent years, and it is in their own and the Society's interests to seek the maximum conciseness and simplification of symbols and equations consistent with clarity.

7. *Revised Manuscripts.*—When papers are submitted in revised form it is especially requested that they be accompanied by the original MSS.

Reading of Papers at Meetings

8. When submitting papers authors are requested to indicate whether they will be willing and able to read the paper at the next or some subsequent meeting, and approximately how long they would like to be allotted for speaking.

9. Postcards giving the programme of each meeting are issued some days before the meeting concerned. Fellows wishing to receive such cards whether for Ordinary Meetings or for the Geophysical Discussions or both should notify the Assistant Secretary.

MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 115 No. 3

MEETING OF 1955 MARCH 11

Professor Sir Harold Jeffreys, President, in the Chair

The election by the Council of the following Fellow was duly confirmed :—

Valdemar Axel Firsoff, M.A., The Retreat, Pilton, Shepton Mallet, Somerset (proposed by P. Moore).

Ninety-three presents were announced as having been received since the last Meeting, including :—

A. C. B. Lovell, *Meteor Astronomy* (presented by the author) ; and
P. A. Moore, *Guide to the Planets* (presented by the author).

MEETING OF 1955 APRIL 6

Professor Sir Harold Jeffreys, President, in the Chair

The election by the Council of the following Fellows was duly confirmed :—

Ian Robin Holton Brickett, 69 Dunvegan Street, Sydenham, Johannesburg, South Africa (proposed by W. H. van den Bos);

George Stephen Brosan, Ph.D., B.Sc.(Eng)., A.M.I.E.E., A.M.I.Prod.E., The Polytechnic, Regent Street, London, W.1 (proposed by P. Moore);

Walter George Grimwood, 21 Southlands Road, Bexhill-on-Sea, Sussex (proposed by L. S. T. Symms);

Hugh M. Johnson, Ph.D., State University of Iowa, Iowa City, Iowa, U.S.A. (proposed by R. H. Garstang); and

Charles Alexander Shain, B.Sc., Radiophysics Laboratory, C.S.I.R.O., University Grounds, City Road, Chippendale, N.S.W., Australia (proposed by J. L. Pawsey).

The President announced that the Council had elected the following an Associate of the Society :—

Donald Howard Menzel, Professor of Astrophysics and Director of the Harvard College Observatory, Cambridge, Massachusetts, U.S.A.

Eighty-one presents were announced as having been received since the last Meeting, including :—

A. Dauvillier, *Le magnétisme des corps célestes*, Volume I (presented by the author); and

G. J. Whitrow, *La Structure de l'Univers*—revised French edition, translated by G. de Vaucouleurs (presented by the author).

Professor Dirk Brouwer delivered the George Darwin Lecture (see p. 221), taking as his subject " The motions of the outer planets ".

After the lecture, the President delivered a short address, and presented the Gold Medal of the Society to Professor Brouwer, to whom it had been awarded for his outstanding contributions to celestial mechanics. The President said :—

" Most of you were present in February when Dr. Jackson explained the Council's reasons for awarding the Gold Medal to Professor Dirk Brouwer. As I was not on the Council at the time, it may be appropriate for me to say something about why I should have supported the award had I been on the Council when it was made. This is perhaps not altogether superfluous, because Council's decisions have not always been unanimously approved. I refer to 1848, when the Council, being apparently unable to choose between Adams and Leverrier, awarded testimonials to them and to ten other gentlemen, mostly for reasons unconnected with the discovery of Neptune. But they did not include Galle, who was the first person to see the planet and know what it was !

" We often hear that dynamical astronomy is a dead subject, which means that it is a hard subject; most of the easy problems have been done. Further, a solution that has fitted for 50 years generally needs revision in 100 years; more terms have to be included and the labour goes up like, roughly, the cube of the number of terms. The understanding of the physical constitution of the bodies of the Solar System needs all the accuracy we can get.

" I first heard of Professor Brouwer when he collaborated with de Sitter on the revision of the fundamental constants, especially those related to the Figure of the Earth. He also did one of the fundamental papers in the theory of the motion of Hyperion. He will be telling you himself about his work on the outer planets. His work on the secular accelerations of the Sun and Moon is of the first importance, though perhaps he finds it as difficult as I do to see what it means.

" Professor Brouwer : I present to you this Medal as an indication of our esteem for your past work, and accompany it with best wishes for all you may undertake in the future."

THE MOTIONS OF THE OUTER PLANETS

George Darwin Lecture, delivered by Professor Dirk Brouwer on 1955 April 6

It has been my good fortune to have been closely associated with three men who received the honour of the award of the Gold Medal of the Royal Astronomical Society: Willem de Sitter, under whom I studied at Leiden; Ernest Brown, whose research assistant I became after I left Leiden; and Frank Schlesinger, my predecessor as director of the Yale Observatory. Two of these men were George Darwin Lecturers. It is with a sense of humility that I appear before you this afternoon to contribute to the series that has produced so many memorable lectures.

I am expected to speak about my own work, which is like inviting the entire astronomical world to scrutinize it and assay it on its merits. This is not to be undertaken lightly, but I am encouraged by a piece of advice that Ernest Brown gave me long ago when I had been asked to perform a task that up to that time had always been the responsibility of more experienced men in the University. He merely said: "Just do as you are told". I have followed this advice many times since in similar situations, and I remembered it again as I prepared for this occasion.

During the second half of the nineteenth century the motions of the principal planets were the subject of extensive investigations: in France by Leverrier, continued by Gaillot; in the United States by Newcomb and Hill. These investigations formed the basis of the tables that are in current use for the construction of the planetary ephemerides published in the national ephemerides. Since these tables were constructed there has been an impressive accumulation of observations, generally of higher accuracy than the observational material that was available for determining the constants of the tables. This alone is a sufficient reason for a resumption of activity in this field at the present time.

A second reason is the rapid development in the field of high-speed calculators in recent years which has opened up possibilities for the study of problems in celestial mechanics that were beyond the dreams of our predecessors of even one generation ago. This applies both to the handling of large bodies of observational data and to theoretical investigations necessary to improve the planetary tables.

Concerning the terrestrial planets the discussions of the observations of the Sun by Morgan and Scott (**1**) and by Spencer Jones (**2**), and of Mercury by Clemence (**3**) are examples of the very considerable increase in our knowledge of these orbits over that contained in the famous summary published by Newcomb (**4**). This work on the terrestrial planets is being continued at the United States Naval Observatory; a discussion of the observations of Venus by Mr R. L. Duncombe is approaching completion, while Dr Clemence has in hand an entirely new theory of the motion of Mars. The first part of this, containing

the perturbations of the first order, has been published (**5**); the remaining portions of the theory have since been completed, and the comparison with observations is far advanced.

A new departure in dealing with the orbits of the outer planets was the "Coordinates of the Five Outer Planets" by Eckert, Brouwer and Clemence (**6**). These orbits were obtained by numerical integration. They were adjusted to represent the existing observations and will provide accurate ephemerides for a century in the future, except that for Pluto no new comparison with observations was made. Pluto's orbit, based on Bower's elements XIX (**7**), may require appreciable corrections to represent the observations of the twenty-first century.

In the numerical integration of these orbits the attractions by the terrestrial planets were ignored because including them would have required a much smaller interval. A supplementary volume by Clemence (**8**) provides the necessary corrections.

This, briefly, is the present condition of the improvements of the orbits of the principal planets. I intend to review a few of the details of the work on the outer planets and indicate in what direction we hope to carry on further research on this problem.

The masses of the outer planets.—For the numerical integration we were content to adopt the same masses for Jupiter, Saturn, Uranus and Neptune as had been used by Newcomb and Hill. This had the advantage of making the numerical integration orbits more directly comparable with the planetary tables. For Pluto we adopted the mass $\frac{1}{360000}$, which had been obtained from a solution based on Wylie's residuals of Neptune.

In the course of further work on the numerical integrations of these planets it is our intention to provide coefficients necessary for applying corrections to the masses of the five planets.

A recent investigation by Hertz (**9**) indicates that the mass of Saturn should be increased by one per cent. From a new discussion of the observations of Jupiter which extends Hill's solution to the present time, he finds for the reciprocal of the mass $3497{\cdot}64 \pm 0{\cdot}27$ (p.e.). This brings the value obtained from planetary perturbations into excellent agreement with that resulting from satellite orbits.

For the mass of Neptune, the most comprehensive determination is that by Eichelberger and Newton from the observations of Triton (**10**). They find a mass in good agreement with Newcomb's value, but there is a large difference between the result from the visual and the photographic observations. In fact, there are good reasons to suspect systematic errors in both. In the years 1939–1942 Alden (**11**) made a series of photographic observations of this satellite with the Yale telescope, then in Johannesburg. An objective grating was used to produce first-order spectral images of the planet of nearly equal magnitude to the image of the satellite. This material is now being used by Miss Jocelyn Gill at the Yale Observatory, with as one ojbect a new determination of the mass of the planet. For this reason special care is given to the scale value.

It is encouraging that other new work in this field is planned or in progress. Dr Daniel Harris has undertaken a study of satellite orbits with the Yerkes refractor, and has done a considerable amount of work on the satellites of Uranus. An extensive series of photographic observations of Saturn's satellites made with the Yale telescope in the southern hemisphere was measured at the United

States Naval Observatory by Mr F. H. Hollander. The reductions have recently been taken up by Dr Morris S. Davis, a former student of mine, now at the University of North Carolina. One of the principal objects of this study is the determination of Saturn's mass. For the mass of Jupiter Dr Clemence has undertaken the preliminary steps toward the realization of a proposal made by Hill (**12**) in 1873. Hill selected 13 minor planets near the 2 : 1 resonance region with Jupiter for which the principal Jupiter perturbations in longitude have coefficients between 1°·4 and 9°, and periods between 60 and 120 years. These planets are being observed at the United States Naval Observatory; the old observations have been collected and are being reduced to a uniform system. Dr Clemence expects that with the aid of electronic calculators it will soon be possible to calculate orbits of the requisite precision.

In the Introduction of the volume of the "Coordinates of the Five Outer Planets" a solution is made for the mass of Pluto from its action on both Uranus and Neptune, based upon the residuals obtained from Wylie's discussion (**13**). Since then Woolard (**14**) has made a new compilation of the pre-discovery observations of Uranus and compared them with the orbit obtained by numerical integration. I have recently made a new solution for Uranus in which the pre-discovery observations are included. The accidental error of these old observations is disappointingly large. The r.m.s. residuals in longitude and latitude are 5″·9 and 4″·1, respectively. The solution depends to a large extent on the residuals of Flamsteed's observations, one in 1712, three in 1715. These are not accordant. Rejection of the observations in 1715, as is favoured by Kourganoff (**15**), yields a mass of Pluto comparable with or greater than the Earth's mass. It appears to me that the only safe procedure is to give all pre-discovery observations equal weight and to refrain from rejecting any. Weights 0·05 and 0·10 in Wylie's system were assigned to the longitude and latitude equations, respectively. The result of the solution, if μ is the fractional correction to the mass of Pluto used in the integration, is

$$\mu = -0{\cdot}58 \pm 0{\cdot}44 \text{ (m.e.)}.$$

A new solution for Neptune from four-year means gives

$$\mu = -0{\cdot}06 \pm 0{\cdot}14 \text{ (m.e.)}.$$

This mean error is 60 per cent greater than in the published solution. This must be ascribed to the use of four-year means rather than the individual normals given by Wylie, an indication of lack of independence of the normals. A weighted mean of these solutions is:

$$\mu = -0{\cdot}10 \pm 0{\cdot}13 \text{ (m.e.)}$$

$$m_{\mathrm{P}} = \frac{1}{400000}\{1 \pm 0{\cdot}10 \text{ (p.e.)}\}.$$

The Neptune residuals look very satisfactory; the longitude solution depends entirely on the Lalande residual of 1795, but the modern latitudes cannot be satisfied with a much smaller mass. The Uranus residuals are less sensitive to the mass of Pluto. The latitudes contribute practically no weight at all to the determination, but, regardless of the adopted mass of Pluto, they show a systematic trend.

Essentially the same systematic trend in the residuals in latitude was found by Newcomb when he compared the observations with his first theory of Uranus (**16**).

Newcomb represented the residuals by an expression which becomes

$$+0''\cdot18-0''\cdot28T$$

if T is expressed in centuries from 1900. He concluded that at the epoch 1840 Auwers' equatorial declinations are too small by $0''\cdot35$ and that this error is diminishing at the rate of $0''\cdot28$ per century. This conclusion seemed reasonable enough in 1870. Wylie obtained the result

$$+0''\cdot13-0''\cdot21T$$

from the observations since 1830, compared with Newcomb's second tables. He states that the declination system of the FK3, to which he reduced the observations, is not likely to be affected by a systematic error of this size. The observations since 1870 compared with the numerical integration orbit give:

$$+0''\cdot12-0''\cdot35T.$$

This persistent unexplained deviation in the motion of Uranus destroys much of the confidence in the solution of Pluto's mass from Neptune. For if the latitudes of Uranus show the presence of an unknown systematic effect, how can we reasonably expect that the latitudes of Neptune are free from a similar unknown effect that is absorbed in the solution for the inclination and node and the mass of Pluto?

If we assume, against better judgement, that the systematic trend in the Uranus residuals indicates an error in the FK3 declinations, we must correct the Neptune residuals in latitude by subtracting the linear term found from Uranus. If this is done, the solution from the latitudes of Neptune gives $\mu=+0\cdot39$, increasing the mass of Pluto.

Kuiper (**17**) objects to a mass of Pluto comparable to the mass of Venus on the grounds of Pluto's apparent faintness and small measured diameter. Until recently, judging by the evidence from Neptune alone, it appeared to me that it would be difficult to bring this physical estimate into agreement with the gravitational determination. I have now come to the conclusion that the gravitational determination is not entitled to the weight that the formal solution indicates. The longitudes of Uranus and Neptune combined give

$$\mu=-0\cdot20\pm0\cdot24\ \text{(m.e.)},$$

$$m_P=\frac{1}{450\,000}\{1\pm0\cdot20\ \text{(p.e.)}\}.$$

Much of the weight depends on the observations of Neptune by Lalande in 1795. But Jackson (**18**) showed in 1930 that a wide range in the elements of Neptune is possible without requiring the presence of improbably large systematic errors in the observations. For a slowly moving object such as Neptune these systematic errors are in part absorbed in the solution and the resulting weight is fictitiously high. Thus, after a great deal of work has been done on the problem, we must for the present leave the question in this state of uncertainty. As Kourganoff remarks, the perturbations by Pluto in the longitude of Uranus will soon increase. By 1965 the difference in longitude between $m_P=0$ and $m_P=\frac{1}{400000}$ will amount to $0''\cdot3$ and the weight of the solution from the Uranus longitudes will have increased fourfold compared with the weight of the solution that included observations to 1938. Similarly, the perturbations on Neptune must eventually increase so that the solution from the longitudes of Neptune will become less dependent on the Lalande residual of 1795.

Only with the passage of time may we expect a reliable gravitational determination of Pluto's mass and a clue to the cause of the puzzling deviation of Uranus in latitude. The latter may well be of an exciting nature.

Accuracy of the numerical integration orbits.—As a test of general interest Clemence and I (**19**) evaluated from the coordinates and velocity components of the numerical integration orbits the three integrals of areas and the energy integral for five dates, two of which are very near the beginning and the end of the integration, while one coincides with the osculation date. The differences from constancy are on the whole in accordance with expectation on the basis of accumulation of error of rounding, though in the case of one of the integrals of areas the discordance among the individual values is approximately six times that expected. We tentatively ascribe this to truncation errors which may arise in Jupiter and Saturn. The effect is of little practical interest since it does not reach the last printed decimal place in the coordinates. Yet some of the details have at least the interest of a curiosity. As stated in the Introduction to the volume, the tenth differences of the accelerations of Jupiter and Saturn show from time to time systematic trends resembling harmonic functions of relatively short period. Further examination showed that they occur in Jupiter near each perihelion, in Saturn near each conjunction with Jupiter, recurring in cycles of about 60 years. An unforeseen circumstance is that the arbitrarily chosen osculation date, 1941 January 6·0 = JD 2430000·5 is very near a Jupiter perihelion and a conjunction of Jupiter and Saturn. It would have been difficult to choose a less suitable date for overlapping the forward and backward integrations.

All the available evidence indicates that for the interval of four centuries for which the numerical integrations were carried out, these orbits are of a quality that would be extremely difficult to attain with general theories. There remains the objection that in the past we have become accustomed to providing planetary ephemerides on the basis of general planetary theories.

It has often been said that the drawback of the numerical integration procedure is that the orbit is determined only for the period for which the integration has been carried out, and that no information is obtained concerning the general properties of the orbit. In this connection I refer to a recent paper by Stiefel (**20**), in which the relation between numerical integrations and general planetary theories is considered. Dealing with a system of the Sun and only two planets, he shows that it is, at least in principle, possible to derive from the result of the numerical integration of the orbits a representation that corresponds to that of a general planetary theory in the ordinary form. Perhaps with powerful electronic computing facilities a method of practical value could be developed. Even then an extension to more than two planets would lead to formidable complications.

Whatever the merits of such an approach may be, it would indeed be regrettable if in the future purely numerical methods were to supplant analytical methods completely. The numerical methods should rather be looked upon as an aid for testing the results of analytical theories and for stimulating further theoretical developments. This sentiment was expressed by Danjon (**21**), and I concur wholeheartedly. I shall now turn to the subject of general planetary theories.

Classical planetary theory in rectangular coordinates.—In the introductory section of Hill's " Researches in the Lunar Theory " (**22**), the author remarks that

in elliptic motion the rectangular coordinates can be developed in periodic series in the mean anomaly each coefficient of which can be stated as a single Bessel function. The law of the series is evident, and the approximation can be carried as far as we wish. The longitude and the latitude have much more complicated expressions; the coefficients in the series cannot be expressed by a finite number of Bessel functions. Hill then continues: "If this is true in the elliptic theory, how much more likely is a similar thing to be true when the complexity of the problem is increased by the consideration of disturbing forces?"

For numerical integration Encke had introduced the differential method in rectangular coordinates, being attracted by the simplicity and symmetry of the equations. He subsequently made an attempt to apply the same principles to the problem of general perturbations (**23**), but in the form developed by him the perturbation in the radius vector appears as a fourth quantity to be obtained in the course of the integration of the equations. The introduction of two superfluous constants of integration is a severe drawback. It is curious to note that both Brünnow, in an appendix to Encke's article, and Hill (**24**) use Encke's equations as stepping stones toward the derivation of the familiar equations in polar coordinates.

There is an indication of a promising approach to obtaining suitable equations in rectangular coordinates in Dziobek's treatise (**25**), but Dziobek was satisfied with a final result in symbolic form which he described as unsuitable for practical use.

The principle is the following: For convenience choose an elliptic reference orbit nearly coinciding with the orbital plane of the perturbed body so that the z-coordinate is of the order of the perturbations. Let x_0, y_0 designate the coordinates in the reference orbit as functions of the time and the chosen fixed elements. Let the perturbed coordinates be $x_0+\xi$, $y_0+\eta$. The equations for ξ and η are

$$\frac{d^2\xi}{dt^2}+\frac{\mu\xi}{r_0^3}-\frac{3\mu x_0}{r_0^5}(x_0\xi+y_0\eta)=X=\left(\frac{\partial R}{\partial x}\right)+\Delta X,$$

$$\frac{d^2\eta}{dt^2}+\frac{\mu\eta}{r_0^3}-\frac{3\mu y_0}{r_0^5}(x_0\xi+y_0\eta)=Y=\left(\frac{\partial R}{\partial y}\right)+\Delta Y.$$

The right-hand members consist of the x and y components of the disturbing force computed with the coordinates of the reference orbit with the additions ΔX, ΔY, which consist of squares and products of ξ and η but may also contain higher-order parts of the disturbing forces.

If the right-hand members are ignored, the equations have the particular solutions

$$\xi_j=\frac{\partial x_0}{\partial c_j},\quad \eta_j=\frac{\partial y_0}{\partial c_j},\quad j=1,2,3,4,$$

in which c_j are any set of four orbital elements that define the motion in the plane of an elliptic orbit. Thus the complete solution of the equations without right-hand members is

$$\xi=\sum_{j=1}^{4}C_j\frac{\partial x_0}{\partial c_j},\quad \eta=\sum_{j=1}^{4}C_j\frac{\partial y_0}{\partial c_j},$$

where C_j are four arbitrary constants. By the method of the variation of arbitrary

constants the same form of solution can be used if C_j are obtained by the integration of dC_j/dt from the equations

$$\left.\begin{aligned} \Sigma \frac{\partial x_0}{\partial c_j}\frac{dC_j}{dt} &= 0, & \Sigma \frac{\partial y_0}{\partial c_j}\frac{dC_j}{dt} &= 0, \\ \Sigma \frac{\partial \dot{x}_0}{\partial c_j}\frac{dC_j}{dt} &= X, & \Sigma \frac{\partial \dot{y}_0}{\partial c_j}\frac{dC_j}{dt} &= Y. \end{aligned}\right\} \tag{1}$$

In a sense this completes the solution, but without further modification the method is hopelessly laborious and has a number of practical disadvantages. The principal drawback is that the resulting expressions require numerous multiplications of series, both in the integrands and outside the integrals. This is overcome (**26**) by using the device due to Hansen of bringing a factor outside an integral sign into the integrand by treating the time it contains as constant in the integration. This requires this t to be kept distinct from the t already present in the integrand, which is accomplished by introducing a different symbol, say, τ for the t so arising. After integration τ is to be restored to t. Functions in the integrand to be expressed in terms of τ are written with a bar over them. The final expressions for the perturbations ξ and η become

$$\xi = \frac{\partial x_0}{\partial l}\delta^n l + \int (A_x X + A_y Y)\,dt, \quad \eta = \frac{\partial y_0}{\partial l}\delta^n l + \int (B_x X + B_y Y)\,dt, \tag{2}$$

in which

$$\begin{aligned} \delta^n l &= -\frac{3n^2 a}{\mu}\iint\left(\frac{\partial x_0}{\partial l}X + \frac{\partial y_0}{\partial l}Y\right)dt\,dt \\ &= -\frac{3n^2 a}{\mu}\iint\left\{\left(\frac{\partial R}{\partial l}\right)_0 + \frac{\partial x_0}{\partial l}\Delta X + \frac{\partial y_0}{\partial l}\Delta Y\right\}dt\,dt. \end{aligned} \tag{3}$$

The functions

$$A_x = (x_0, \bar{x}_0), \quad A_y = (y_0, \bar{x}_0), \quad B_x = (x_0, \bar{y}_0), \quad B_y = (y_0, \bar{y}_0) \tag{4}$$

are written as Poisson brackets because this is the form in which they are expressed in terms of partial derivatives of x_0, y_0 with respect to a set of elliptic canonical variables. They are of the same nature as A and B which arise in Hansen's method in the function W from which the Hansen perturbations $n\delta z$ and ν are obtained. It is unnecessary to deal with the perturbations perpendicular to the plane of the reference orbit. It is evident that the expressions (2) and (3) can at once be generalized for any arbitrary fixed coordinate system. With the particular choice of the xy-plane the equation for z reduces to a form common to that in other types of planetary theory.

Some similarity with details of Hansen's method is evident; yet the final expressions are totally different. A comparison between the two methods involves much more than a mere counting of the number of operations required. An actual application of the method was made by Davis (**27**), who also made a parallel computation by Hansen's method. A similar comparison was made by Clemence (**28**), who also comments in part in the Celestial Mechanics section of the Reports on the progress of astronomy in 1951 (**29**): "The calculation of the perturbations of higher order than the first is greatly simplified. In addition to this important advantage, the developments possess an unusual degree of symmetry. This symmetry is obtained at some sacrifice of brevity in the numerical development of these factors, which renders Hansen's method the readier one if it is proposed to stop with the first approximation, and which to

some extent counterbalances the advantages of the new method in the second approximation. There is no doubt, however, that in the third and higher approximations the new method is preferable to all others. Its use would be indicated if new theories of the outer planets were to be constructed and, in most cases where it is applicable, if it were desired to obtain a general theory of a minor planet comparable in precision with the observations."

This, I believe, is sufficient justification for the use of the method in rectangular coordinates for the development of new theories of the planets Uranus, Neptune, and Pluto in which the perturbations of higher order are important. This work is now in progress at the Yale Observatory as a joint undertaking by my colleague, Dr van Woerkom, and myself.

For the development of the odd negative powers of the distance, the method given a few years ago by Clemence and myself (**30**) proved useful. It assumes that an approximate development of the reciprocal of the distance is available. Denote this by Δ^{-1}. Let the square of the distance be D^2, which can be computed as a double Fourier series with all the accuracy desired. The square Δ^{-2} and the product $D^2\Delta^{-2} = 1 + \epsilon$ are then obtained. The deviation ϵ from unity shows the error in Δ^{-1}. It is then easy to obtain

$$D^{-2} = \Delta^{-2}(1 - \epsilon + \epsilon^2 - \epsilon^3 + \dots),$$
$$D^{-1} = \Delta^{-1}(1 - \tfrac{1}{2}\epsilon + \tfrac{3}{8}\epsilon^2 - \tfrac{5}{16}\epsilon^3 + \dots),$$

and from these by multiplication of series all the odd negative powers of D required. With adequate automatic computing equipment the method is probably not more elaborate than the conventional development. All of the operations can be performed in strictly routine fashion, and the operator has to learn only one unfamiliar technique, the multiplication of series. The simplicity and certainty of the method are its principal advantages.

For applications to the orbits of the principal planets provisional developments of the reciprocals of the distances in most cases are available. This is not so for the distances between Pluto and the four major planets. For these, provisional developments were obtained by harmonic analysis. Only the distance Neptune–Pluto gave serious difficulties, but harmonic analysis from 80 special values gave an expansion that could serve as a first approximation.

The work now has the highest priority on our computing programme. It is a task of which my colleague Dr van Woerkom has taken an important share. In fact, all of the machine calculations are performed either by him or under his immediate supervision.

Generalized planetary theory.—Before the decision was reached to concentrate on this problem, a considerable amount of experimentation on a more ambitious project had been carried out. The usual form adopted for general planetary theories in which powers of the time are permitted in the coefficients of periodic terms is imposed on us on account of the difficulty of proceeding otherwise. Hill once called this form a "temporary makeshift". A much more satisfying solution would be one like that of the main problem of the lunar theory in which linear combinations of the four fundamental arguments, D, l, l', F appear in the arguments of the periodic terms of the Moon's longitude, parallax and latitude, and no powers of t appear in the coefficients. In the satellite systems the mere fact that the motions of the pericentres and nodes proceed so much more rapidly than in the planetary system compels us to adopt this form.

The literature of the second half of the nineteenth century is rich in investigations on this form of solution for the planetary problem. Most of these were qualitative in character, but Gyldén and Hill were especially concerned with devising methods that could be used to obtain a representation of the orbits of the principal planets. In the fourth volume of his "Collected Mathematical Works", Hill returns to the problem again and again. He devoted particular attention to the application of Delaunay's method to the planetary problem. It is perhaps curious that Hill never conceived the modification of Delaunay's method in which the non-critical terms in the disturbing function are treated together in one operation. This approach was elaborated in much detail by v. Zeipel (**31**) and also treated by Brown and Shook (**32**). Unfortunately, the procedure requires that all the parameters involved are left indeterminate in the formulas. As a consequence, the degree of complexity is forbidding. Thus, while the problem has been foremost in the minds of some of the outstanding men in the field, we still lack a satisfactory representation of planetary motions valid for long intervals of time. This is not to be ascribed to lack of competence or perseverance of these investigators but merely to the magnitude of the problem.

The only completed projects along this line are solutions of the equations of the secular variations by Leverrier, Stockwell, Harzer and recently with improved values of the planetary masses by van Woerkom and myself (**33**). The scope of this solution is extremely limited compared with that of providing fully general theories of planetary motion. Yet it has at least the merit of indicating the form of a general analytical perturbation theory in which the time is present only in the arguments of sine and cosine terms.

It seems worth while to explore the possibility of using the solution of the secular variations as the basic data for a more general attack on the problem. This indicates a representation in which the arguments of the trigonometric functions appear in a literal form, while the coefficients are to be treated numerically. An important decision to be made is whether to aim at providing for all the principal planets at once, or begin with a more modest undertaking.

There is a good reason for concentrating first on the outer planets because, though they produce large secular variations in the elements of the terrestrial planets, the effects of the terrestrial planets on the outer planets are insignificant. It would not be possible at present to extend the solution to the planets Uranus, Neptune and Pluto. The orbits of Neptune and Pluto are such that if the longitudes of the perihelia and nodes of these planets are permitted to vary without restriction, the orbits will at some time intersect. This introduces a singularity into the problem that makes it impossible to develop the disturbing function in the general form that is required. For this reason, van Woerkom and I were obliged to leave Pluto out of the solution of the secular variations. This does not necessarily indicate that no solution of the desired form exists. The three outermost planets have mean motions so nearly in the ratio 6 : 3 : 2 that it is not excluded that their orbits are actually connected by a resonance relationship. The completion of the ordinary planetary theory for these planets may help to decide whether this conjecture has any merit. At the present time, therefore, we have no choice but to begin with the system consisting of the Sun, Jupiter and Saturn only. Further, to simplify the problem, we will ignore the inclinations. This surely is the simplest case that, if solved, could serve as the nucleus of a more ambitious project.

Let λ be the mean longitude, a the semi-major axis, e the eccentricity and ϖ the longitude of the perihelion, ψ the true longitude, r the radius vector, and put

$$\psi = \lambda + \epsilon, \quad r = a(1+\rho).$$

Then

$$x = a(1+\rho)\cos(\lambda+\epsilon), \quad y = a(1+\rho)\sin(\lambda+\epsilon).$$

If

$$1+\xi = (1+\rho)\cos\epsilon, \quad \eta = (1+\rho)\sin\epsilon,$$

then

$$x = a(1+\xi)\cos\lambda - a\eta\sin\lambda,$$
$$y = a(1+\xi)\sin\lambda + a\eta\cos\lambda.$$

In these expressions, a and λ appear explicitly; ξ and η are expressible as cosine or sine series the arguments of which are multiples of the mean anomaly, $l = \lambda - \varpi$, while the coefficients are power series in the eccentricity. If these expressions apply to Jupiter, similar expressions with primed quantities could be written for Saturn.

In the usual form of planetary theory the disturbing function is developed by using constants for a, e, ϖ and for λ a linear function of the time. In a first development of a generalized theory it is necessary in the first place to introduce for e and ϖ the expressions that are obtained in the solution of the secular variations. In addition it is possible to incorporate the principal perturbations in the mean longitude, transformed to the desired form.

The principal terms in Hill's solution for the eccentricities and perihelia of Jupiter and Saturn (**34**) are

$$\left.\begin{aligned} e\begin{matrix}\sin\\ \cos\end{matrix}\varpi &= +0{\cdot}043749\begin{matrix}\sin\\ \cos\end{matrix}\chi + 0{\cdot}015683\begin{matrix}\sin\\ \cos\end{matrix}\chi', \\ e'\begin{matrix}\sin\\ \cos\end{matrix}\varpi' &= +0{\cdot}032735\begin{matrix}\sin\\ \cos\end{matrix}\chi - 0{\cdot}048268\begin{matrix}\sin\\ \cos\end{matrix}\chi', \\ \chi &= +4''{\cdot}0995t + \text{const.}, \\ \chi' &= +27''{\cdot}7301t + \text{const.}, \end{aligned}\right\} \qquad (5)$$

t being expressed in years. The designations of χ and χ' have been interchanged compared with Hill. Substituted into the expressions for ξ and η, if λ_0 and λ_0' are linear functions of t, the results are:

Jupiter

Argument	ξ cosine series	η sine series
$\lambda_0 - \chi$	−0·04379	+0·08746
$\lambda_0 - \chi'$	−0·01571	+0·03134
$2\lambda_0 - 2\chi$	+0·00095	+0·00048
$2\lambda_0 - \chi - \chi'$	+0·00068	+0·00034
$2\lambda_0 - 2\chi'$	+0·00010	+0·00005
$\chi - \chi'$	−0·00034	...
$2\chi - 2\chi'$	+0·00001	...

Saturn

Argument	ξ' cosine series	η' sine series
$\lambda_0' - \chi$	−0·03281	+0·06540
$\lambda_0' - \chi'$	+0·04835	−0·09645
$2\lambda_0' - 2\chi$	+0·00053	+0·00026
$2\lambda_0' - \chi - \chi'$	−0·00157	−0·00078
$2\lambda_0' - 2\chi'$	+0·00122	+0·00061
$\chi - \chi'$	+0·00083	...
$2\chi - 2\chi'$	−0·00003	...

The expressions (5) are then used to transform Leverrier's results for the great inequalities in the mean longitudes of Jupiter and Saturn. They are:

Long-period terms in the mean longitude in radians

Argument	(sine series) Jupiter	Saturn
$2\lambda_0-5\lambda_0'-\chi+6\chi'$	−0·00000	+0·00001
$2\lambda_0-5\lambda_0'-2\chi+5\chi'$	+0·00005	−0·00013
$2\lambda_0-5\lambda_0'-3\chi+4\chi'$	−0·00062	−0·00153
$2\lambda_0-5\lambda_0'-4\chi+3\chi'$	+0·00531	−0·01311
$2\lambda_0-5\lambda_0'-5\chi+2\chi'$	−0·00167	+0·00411
$2\lambda_0-5\lambda_0'-6\chi+\chi'$	+0·00011	−0·00027
$2\lambda_0-5\lambda_0'-7\chi$	−0·00001	−0·00002

It would have been possible to transform all the larger terms of Leverrier's theory into the required form, but it seemed doubtful that much could be gained in this way in the first approximation. The only additional terms actually used are the synodic perturbations, which are comparable in magnitude to the quadratic terms in the equation of the centre. It should be noted that the principal coefficients of the great inequalities in the mean longitudes are intermediate in magnitude between the coefficients in the equation of the centre that correspond to the first and the second powers in the eccentricities. Hence, in a numerical development the inclusion of these terms is bound to accelerate the approximations.

The principal task in the development of the disturbing forces is the expansion of the odd negative powers of the mutual distance. In the particular problem before us, the inverse cube of the distance, or rather $a'^3\Delta^{-3}$, is the basic function, a' being a constant nearly equal to the mean value of Saturn's semi-major axis.

The computation of $a'^{-2}\Delta^2$ as a cosine series with arguments

$$j(\lambda_0-\chi)+j'(\lambda_0'-\chi')+k(\chi'-\chi)$$

is readily obtained. This may be written

$$a'^{-2}\Delta^2=\Delta_0^2+Q,$$

where

$$\Delta_0^2=1+\alpha^2-2\alpha\cos(\lambda_0-\lambda_0'),$$

and

$$\alpha=0{\cdot}545199.$$

The principal terms in Q are

$$+0{\cdot}13209\cos(\lambda_0'-\chi')-0{\cdot}08816\cos(\lambda_0-\chi')$$
$$+0{\cdot}02751\cos(\lambda_0-\chi)+0{\cdot}02632\cos(\lambda_0-2\lambda_0'+\chi').$$

The expansion of $a'^3\Delta^{-3}$ in powers of Q by the binomial expansion

$$\frac{a'^3}{\Delta^3}=\frac{1}{\Delta_0^3}-\frac{3}{2}\frac{Q}{\Delta_0^5}+\frac{15}{8}\frac{Q^2}{\Delta_0^7}-\ldots$$

would have led to an exceedingly long process. This was shortened by the following device. From

$$\Delta_0^{-2}=\frac{1}{1-\alpha^2}\{1+2\Sigma\alpha^p\cos p(\lambda_0-\lambda_0')\}$$

we obtain by multiplication of series

$$K_1=\Delta_0^{-2}Q.$$

Then

$$a'^{-2}\Delta^2 = \Delta_0^2(1 + K_1).$$

Let the principal periodic term in K_1 be $-2\alpha_1 \cos\phi_1$, where α_1 may be positive or negative. Put

$$\begin{aligned} 1 + K_1 &= 1 + \alpha_1^2 - 2\alpha_1 \cos\phi_1 + Q_1 \\ &= \Delta_1^2 + Q_1 \\ &= \Delta_1^2(1 + Q_1\Delta_1^{-2}) \\ &= \Delta_1^2(1 + K_2). \end{aligned}$$

Next, the largest periodic term in K_2 is selected and the process is continued. Thus, finally,

$$\frac{\Delta^2}{a'^2} = \Delta_0^2 . \Delta_1^2 . \Delta_2^2 \ldots . \Delta_n^2(1 + K_{n+1}),$$

$$\frac{a'^3}{\Delta^3} = \Delta_0^{-3} . \Delta_1^{-3} . \Delta_2^{-3} \ldots . \Delta_n^{-3}(1 + K_{n+1})^{-3/2}.$$

Each of the factors in the product $\Delta_0^{-3} . \Delta_1^{-3} \ldots .$ is readily expressed with the aid of the well-known Laplacian coefficients. Only the series for Δ_0^{-3} is slowly convergent, the remaining factors consist of few terms. In principle, this procedure could be continued indefinitely by including in the successive Δ's smaller and smaller terms. Actually we stopped with K_4 and the expansion of $(1 + K_4)^{-3/2}$ was obtained by the binomial theorem. In K_4 there were three terms with coefficients exceeding $\pm 0{\cdot}02$. The fifth power of K_4 was sufficient to assure the accuracy required. In this manner the development of the derivatives of the disturbing function consists of multiplications of harmonic series, all with numerical coefficients. Actually exponentials were used by introducing

$$u = \xi + i\eta, \quad s = \xi - i\eta.$$

Except for the preparatory stages all the calculations were carried out at the Watson Scientific Computing Laboratory. I am indebted to Dr W. J. Eckert for making available the facilities of the Laboratory and to Miss R. B. Jones of his staff for supervizing the calculations.

An indication of magnitude of the operations may be obtained from the following facts: the basic series containing the equations of the centre and the perturbations had been limited to 25 terms for Jupiter, 33 terms for Saturn, rounded to five places of decimals. The development of $a'^3\Delta^{-3}$ was carried to seven places of decimals. The numbers of different arguments of cosine terms present in the principal functions are:

$\Delta_0^{-3} . \Delta_1^{-3} . \Delta_2^{-3} . \Delta_3^{-3}$	248 arguments
$(1 + K_4)^{-3/2}$	1296 arguments
$a'^3\Delta^{-3}$	2198 arguments.

Hill's series for $a'^3\Delta^{-3}$ has 183 arguments, but in his development both sines and cosines are present. The difference is, of course, that in the ordinary general planetary theory the arguments are linear combinations of the two mean anomalies. In the more general development in this case, a third argument, $\chi' - \chi$, is present.

Thus, in $a'^3\Delta^{-3}$, Hill's terms

$+0{\cdot}00912144 \cos(2l-5l')$ $+0{\cdot}48984397 \cos(l-2l')$
$+0{\cdot}01926179 \sin(2l-5l')$ $+0{\cdot}52282903 \sin(l-2l')$

correspond to:

$+0{\cdot}0000010 \cos(2\lambda_0-5\lambda_0'+4\chi-\chi')$ $-0{\cdot}0000188 \cos(\lambda_0-2\lambda_0'+3\chi-2\chi')$
$-0{\cdot}0000002 \cos(2\lambda_0-5\lambda_0'+3\chi)$ $+0{\cdot}0018606 \cos(\lambda_0-2\lambda_0'+2\chi-\chi')$
$-0{\cdot}0004476 \cos(2\lambda_0-5\lambda_0'+2\chi+\chi')$ $+0{\cdot}0589256 \cos(\lambda_0-2\lambda_0'+\chi)$
$+0{\cdot}0062662 \cos(2\lambda_0-5\lambda_0'+\chi+2\chi')$ $-0{\cdot}7153388 \cos(\lambda_0-2\lambda_0'+\chi')$
$-0{\cdot}0214232 \cos(2\lambda_0-5\lambda_0'+3\chi')$ $+0{\cdot}0294726 \cos(\lambda_0-2\lambda_0'-\chi+2\chi')$
$+0{\cdot}0025462 \cos(2\lambda_0-5\lambda_0'-\chi+4\chi')$ $-0{\cdot}0014666 \cos(\lambda_0-2\lambda_0'-2\chi+3\chi')$
$-0{\cdot}0002788 \cos(2\lambda_0-5\lambda_0'-2\chi+5\chi')$ $-0{\cdot}0000084 \cos(\lambda_0-2\lambda_0'-3\chi+4\chi')$
$+0{\cdot}0000154 \cos(2\lambda_0-5\lambda_0'-3\chi+6\chi')$

These samples are representative; for most of Hill's arguments we find seven or eight different arguments. This same multiplication of arguments is found in Leverrier's theory, but Leverrier's longitudes of the perihelia have the instantaneous values.

That the ratio 2198/183 is considerably greater than 7 or 8 must be ascribed to our retention of all coefficients of cosine terms numerically equal to or greater than 2×10^{-7}. Many of the very small coefficients are associated with high multiples of λ_0 and λ_0', and could be safely ignored. As to the number of decimal places, Hill begins with eight for the lower multiples, which number is reduced for the higher multiples to seven, six and finally to five. There is a further advantage on Hill's side: the principal long-period terms in the mean longitudes are obtained in his theory from the development of $a'\Delta^{-1}$, for which he uses eight decimal places in most of the coefficients, but ten for the argument $2l-5l'$. In the more general theory, use could be made of $a'\Delta^{-1}$ at the cost of some further complications in the development. For the present the disturbing forces have been obtained with the development of $a'^3\Delta^{-3}$ only. It would therefore be necessary to carry further developments to additional decimal places if numerical accuracy comparable to Hill's tables were desired. For the immediate future I will consider our goal reached if it turns out that the procedure converges satisfactorily, even if the result cannot compete in the number of significant figures with an ordinary planetary theory.

So far I have dealt only with the development of the disturbing forces. What method of integration is to be used?

There are tremendous advantages in the use of a method in which the coordinates are obtained directly by the integration, especially since it will be necessary to substitute the results of the first integration into the developments for obtaining a second approximation. By using coordinate systems rotating uniformly with the mean mean-motions of the planets, equations in rectangular coordinates are obtained which resemble closely those used by Hill for the lunar theory. The one term in the differential equations of the lunar theory that, as Hill remarks, is a hindrance to their ready integration is the term having the factor r^{-3} not multiplied by the disturbing mass. In the lunar theory this singularity at zero is eliminated with the aid of the Jacobian integral. In the planetary problem with two planets there are terms not multiplied by the perturbing mass with factors r^{-3} and r'^{-3} to contend with. The energy integral contains a linear combination of r^{-1} and r'^{-1}. Obviously there is no possibility of accomplishing the elimination of the singularities at zero. After much experimentation the use of these equations was given up.

A promising alternative approach is to express the derivatives of the elliptic elements in terms of the partial derivatives of the disturbing function with respect to the coordinates. I am now reconciled to the necessity of obtaining perturbations in the elements and having to perform the intermediate operations to secure the coordinates which must be substituted in the disturbing forces for a new approximation. This changing from elements to coordinates adds considerably to the number of multiplications of series that is required.

The crucial question is whether the solution of equations for the eccentricities and perihelia will give unforeseen difficulties. The indications from the work that has so far been done on the problem are favourable. The equations for the semi-major axes and mean longitudes are straightforward. The principal difficulties are those of the small divisors, present in any form of planetary theory. In the Jupiter–Saturn problem the small divisors associated with low multiples of the mean longitudes cause a loss in the number of decimal places after integration. In another respect these are helpful since they prevent the occurrence of other small divisors, except those associated with such very high multiples of the mean longitudes that they are of no astronomical significance. Of course, they are of importance in connection with questions concerning strictly mathematical convergence.

When the developments of the inverse cube of the distance were made, about six years ago, the computations taxed the capacity of the calculators then available. In the meantime new calculators have been developed that are much more efficient for this type of computation. I am now encouraged to proceed with further work on this problem.

I may seem bold to devote so much of this lecture to a project that is still far from completed. The only excuse I can offer is that, unless I had limited myself to generalities, there seemed to be no other way to explain the possibilities of research in gravitational astronomy with modern high-speed calculators.

I have tried to stress the importance of team work in all these undertakings. About eight years ago the Office of Naval Research in Washington decided to lend its support to a research project on the motions of the principal planets, which was then formulated as a joint undertaking of the United States Naval Observatory, the Watson Scientific Computing Laboratory and the Yale University Observatory. If we have been able to accomplish some worthwhile things, it has been possible only with the magnificent support that we have received and with the unselfish cooperation of everyone connected with the work. With continued support and continued team work I am hopeful that we will succeed in completing the unfinished parts of the project.

References

(1) H. R. Morgan and F. P. Scott, *A.J.*, **47**, 193, 1939.
(2) H. Spencer Jones, *M.N.*, **99**, 541, 1939.
(3) G. M. Clemence, *Astr. Papers Amer. Ephem.*, **11**, Pt. I, 1943.
(4) S. Newcomb, *Fundamental Constants of Astronomy, Supp. Amer. Ephem. and Naut. Alm. for 1897*.
(5) G. M. Clemence, *Astr. Papers Amer. Ephem.*, **11**, Pt. II, 1949.
(6) W. J. Eckert, D. Brouwer and G. M. Clemence, *Astr. Papers Amer. Ephem.*, **12**, 1951.
(7) E. C. Bower, *Lick Obs. Bull.*, **15**, 171, 1931.
(8) G. M. Clemence, *Astr. Papers Amer. Ephem.*, **13**, Pt. V, 1954.
(9) H. G. Hertz, *Astr. Papers Amer. Ephem.*, **15**, Pt. II, 1953.

(10) W. S. Eichelberger and A. Newton, *Astr. Papers Amer. Ephem.*, **9**, Pt. III, 1926.
(11) H. L. Alden, *A.J.*, **49**, 70, 1940; **50**, 110, 1943.
(12) G. W. Hill, *Coll. Works*, **1**, 105; *Mem. Amer. Acad. Arts and Sci.*, N.S. **9**, 417, 1873.
(13) L. R. Wylie, *Pub. U.S. Naval Obs.*, Ser. II, **15**, Pt. I, 1942; Pt. III, 1947.
(14) E. W. Woolard, *A.J.*, **57**, 35, 1952.
(15) V. Kourganoff, *Bull. Astr.*, Ser. II, **12**, 147, 1941.
(16) S. Newcomb, *Smiths. Contr. to Knowledge*, No. 262, 176, 1873.
(17) G. P. Kuiper, *Pub. Astr. Soc. Pac.*, **62**, 133, 1950.
(18) J. Jackson, *M.N.*, **90**, 728, 1930.
(19) G. M. Clemence and D. Brouwer, *A.J.*, **60**, 118, 1955.
(20) E. Stiefel, *Arch. Math.*, **5**, 347, 1954.
(21) A. Danjon, *Annu. Bur. Longit.*, 1952, p. A1.
(22) G. W. Hill, *Coll. Works*, **1**, 284; *Amer. J. Math.*, **1**, 5, 1878.
(23) J. F. Encke, *Berl. Astr. Jb. f. 1857*, 319, 1854.
(24) G. W. Hill, *Coll. Works*, **1**, 151; *A.N.*, **83**, 209, 1894.
(25) O. Dziobek, *Die Mathematischen Theorien der Planetenbewegungen*, p. 160, Leipzig, 1888.
(26) D. Brouwer, *A.J.*, **51**, 37, 1944.
(27) M. S. Davis, *A.J.*, **56**, 188, 1952.
(28) G. M. Clemence, *A.J.*, **52**, 122, 1947.
(29) G. M. Clemence, *M.N.*, **111**, 225, 1951.
(30) D. Brouwer and G. M. Clemence, *A.J.*, **52**, 64, 1946.
(31) H. v. Zeipel, *Ark. Mat. Astr. o. Fys.*, **11**, 1, 1916.
(32) E. W. Brown and C. A. Shook, *Planetary Theory*, Ch. VI, Cambridge, 1930.
(33) D. Brouwer and A. J. J. van Woerkom, *Astr. Papers Amer. Ephem.*, **13**, Pt. II, 1950.
(34) G. W. Hill, *Coll. Works*, **4**, 123; *A.J.*, **17**, 81, 1897.

RR TELESCOPII IN 1953 AND 1954, AND THE DEVELOPMENT OF SLOW NOVAE

A. D. Thackeray

(Received 1955 February 10)

Summary

Recent developments in the emission spectrum of RR Tel include the emergence of C IV, [A IV], [K IV], [Fe IV], [Fe V]. A special search for the hitherto unknown [Fe IV] lines proved successful in the light of Edlén's independent predictions from laboratory data. Fe II lines have faded steadily relative to [Fe II].

The rate of development towards higher stages of ionization is compared with that of other slow novae. Time intervals between corresponding stages of RR Pic, RR Tel, RT Ser and η Car are found to be approximately in the ratios 1 : 6 : 17 : 220. One result of this comparison is to give some grounds for supposing that η Car brightened from a pre-nova state not very long before Halley's observation in 1677.

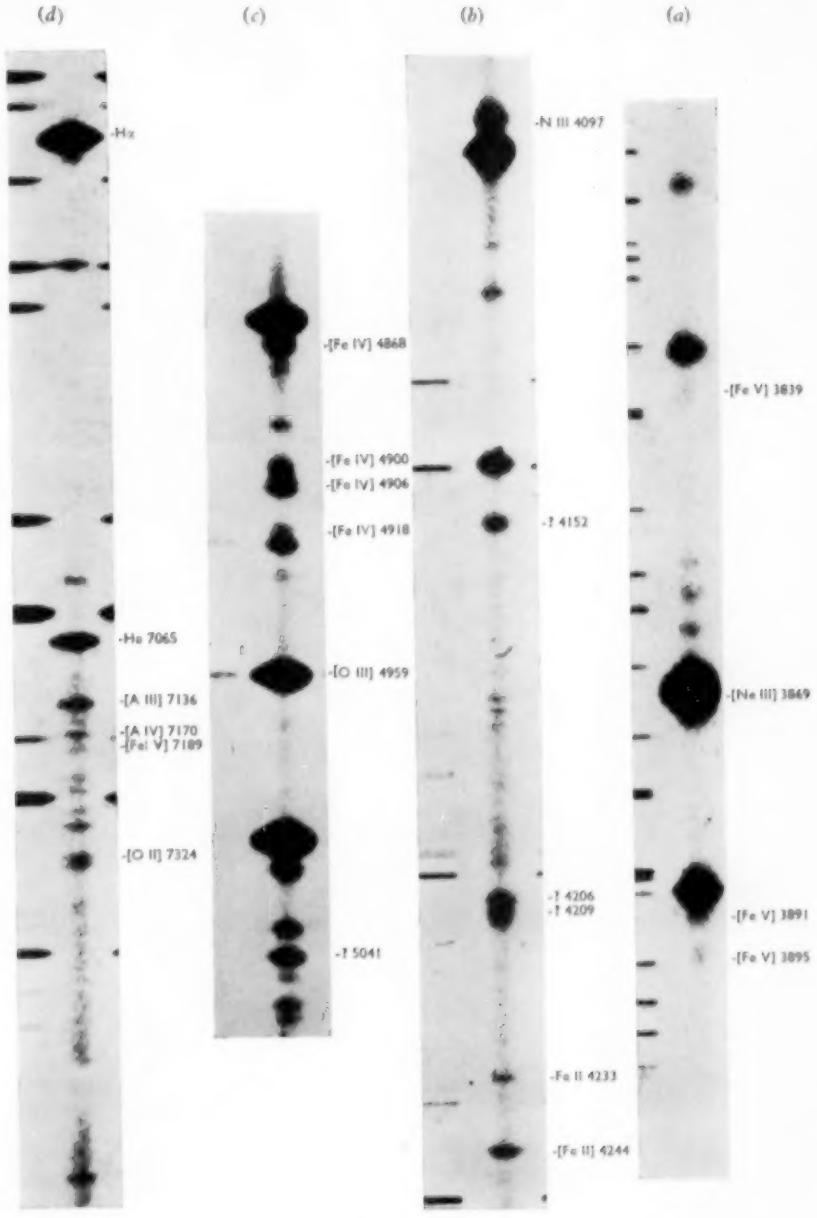
---

1. Radcliffe spectra of the interesting slow nova RR Tel, covering the period 1949–1952, have been described in two previous communications.* It will be recalled that the star, which had been known since 1908 to be variable ($12^m{\cdot}5$ to 16^m), rose sharply to 7th magnitude in late 1944, remaining fairly constant until 1949, since when a gradual decline to about 9th magnitude has set in. In 1949 June to August a fairly rapid transition was recorded on low-dispersion Radcliffe spectra (Paper I) from absorption (H, Ti II) to emission (H, Fe II). The much more detailed Cassegrain spectra of 1951–52 showed an emission spectrum of advanced Eta Carinae type (Paper II); between 1951 August and 1952 May there was a marked strengthening of broad bands due to [O III] and [Ne III].

The present paper is concerned with identifications of lines in spectra of RR Tel obtained in 1953 and 1954 with the two-prism Cassegrain spectrograph attached to the Radcliffe reflector. In addition, one useful spectrum (120 A/mm) extending beyond the Balmer limit to 3300 A was secured in 1954 with the Newtonian spectrograph, which, equipped with an aluminized plane grating, is now freed from all glass absorption.

Table I lists the observational material secured during the period. Not all the plates have been measured, but if the plate was used in the present study the last column of the table shows the region covered by the measures. Dr D. Koelbloed of the Amsterdam Institute obtained some further material in 1954 for the purpose of a spectrophotometric study which will give valuable quantitative information concerning the physical state of the object. One of his plates

* A. D. Thackeray, *M.N.*, **110**, 45, 1950; **113**, 211, 1953. These papers are hereafter referred to as Papers I, II respectively.



RR Telescopii in 1954.
(*a*), (*b*), (*c*), *July* 26.
(*d*) *July* 3.

(taken for the infra-red region) is included in Table I, and the writer is deeply indebted to Dr Koelbloed for the use of this plate. The other plates in the table, and all the measures, are by the writer.

TABLE I

Observations of RR Tel 1953–54

Plate	Date	Emulsion	Exposure (min)	Region measured (A)
c 1341	1953 March 20	103aO	32	
c 1342	1953 March 20	103aF	50	
c 1457	1953 May 13	103aF	80	4630–6793
c 1547	1953 June 24	103aO	60	3869–5042
c 1579	1953 July 2	103aF	270	4630–7065
c 1599	1953 July 20	I N	270	6563–8750
c 1649	1953 Aug. 15	103aF	90	
b 1663	1953 Aug. 22	103aO	120	
b 1719	1953 Sept. 22	103aF	136	4026–6793
c 1898	1954 March 28	103aB	30	
c 1915	1954 April 5	103aB	72	4529–6371
c 1955	1954 May 10	103aB	63	4861–6371
c 2023	1954 June 16	103aB	1	
c 2023	1954 June 16		3	
c 2023	1954 June 16		10	
c 2023	1954 June 16		180	3704–6371
c 2040	1954 July 3	II N	240	6563–8860
c 2040	1954 July 3		12	
c 2041	1954 July 3	103 U	20	
b 2106	1954 July 26	103aB	260	3691–6371
c 2165*	1954 Sept. 8	103 U	124	6678–7378
c 2196	1954 Oct. 5	103aB	70	
G T 2†	1954 Oct. 24	R 55	120	3343–4101

* Observation by D. Koelbloed.
† Newtonian spectrogram.

2. Since the majority of the 250 lines observed in 1953–54 appear already in Table V of Paper II, no useful purpose is served by giving anew a complete table of wave-lengths and identifications. Table II therefore presents a selected list of measured lines, which omits, for instance, the large number of familiar lines due to [Fe II] and Fe II. The criteria for inclusion in Table II have been:

(i) Unidentified lines.

(ii) Lines not observed prior to 1953.

(iii) Lines which have clearly strengthened since 1952.

(iv) A few representative lines, previously observed, to serve as an indication of the intensity-scale; e.g. the Balmer lines in the ultra-violet, [O II] 7324, [O III] 4959, etc.

The arrangement of the table is as in Paper II, an asterisk implying a footnote at the end of the table. The identifications are as before based primarily on Mrs Moore-Sitterly's *Revised Multiplet Table*. In addition, the writer is deeply indebted to Professor B. Edlén for communicating, prior to publication, information concerning wave-lengths of permitted lines due to C, N, O in various stages of ionization, and also to Dr R. Velasco for similar communication concerning Ni III. Professor Edlén has also very kindly communicated information about [Fe IV] (see below).

TABLE II

Selected lines in RR Tel, 1953 *and* 1954

Wave-length (A)	1953 Int.	n	1954 Int.	n	Identification					
3343			3*s*	1	Ne III	2*F*	42·9			
3478			4*N*	1	N IV	1	78·7			
3490:			2*N*	1	N IV	1	83·0, 84·9			*
3632·6			?1*n*	1						
47·0			?1*n*	1						
91·5			0·5	2	H		91·56			
3697·0			1*n*	2	H		97·15			
3703·8			3*n*	2	H		03·85			
11·9			3*n*	2	H		11·97			
14·8			1*n*	1						
21·8			6	2	H		21·94,	S III	2*F* 21·1	*
34·4			4	2	H		34·37			
50·1			3	1	H		50·15			
54·8			1*n*	1	?Fe V	3*F*	55·5			
57·6			1·5*n*	2	?O III	2	59·9, 57·2			
70·6			6*n*	2	H		70·60			
81·8:			?0·5	1	?Fe V	3*F*	82·9			
3797·9			8	2	H		97·90,	S III	2*F* 96·7	
3819·7			5*n*	2	He		19·6,	Fe V	3*F* 20·2	
35·4			12	2	H		35·39			
39·6			1·5*N*	1	Fe V	3*F*	38·9,	Fe V	1*F* 38·1	
59·31			3*n*	2				He II	58·07	
91·65:			2*N*:	2	Fe V	3*F*	91·8			*
3895·75			2*N*	2	Fe V	1*F*	95·7			
3954·4			1	1	?Fe III	120	54·38			*
4071·84			4*N*	2	Fe V	1*F*	71·5			
4097·3	4*n*	2	12	1	N III	1	97·31			*
4152·46	4	2	8*n*	2						*
4181·3			1·5*n*	1	Fe V	1*F*	81·3			
4206·62:	6	2	10*n*	2						*
09·04:	4	2	7*n*	2						*
33·23	3	2	3	2	Fe II	27	33·17			
4244·15	8	2	10	2	Fe II	21*F*	43·98, 44·81			
4379·08			2?*d*	2	N III	17	79·09			
4510·3:			3*N*	2	K IV	2*F*	11·0			
29·55	1*N*	2	2	2	?Mn IV	2*F*	28·3			
4571·24	3	2	4	3						*
4634·37	4	4	10	3	N III	2	34·16			
40·89	10	4	18	3	N III	2	40·6, 41·9			
49·46	1·5*n*	4	3*n*	3	C III	1	49 (blend)			
4658·22	9	4	10	3	Fe III	3*F*	58·1			
4713·32	10	4	12	3	A IV	1*F*	11·4	He	12 13·2	
4740·50	4	4	6	4	A IV	1*F*	40·3			
4861·4	70	4	70	4	H		61·3			
4868·2:	4:	4	8:	4	Fe IV	*F*				*
4900·2	2	4	8	4	Fe IV	*F*				
06·2	6	4	12	4	Fe IV	*F*				
18·1	?1	3	4	4	Fe IV	*F*				
30·8	1*n*	3	2*n*	4	?Fe III	1*F*	30·5			*
4958·79	40	4	50	4	O III	1*F*	58·91			
5033·27	4	4	10	4						
41·55	8	4	15	4				Si II	5 41·1	
56·2	3	3	6	4	Si II	5	56·1			
5060·3	1·5	2	4	4	Fe III	1*F*	60·3			

TABLE II (*continued*)

Wave-length (A)	1953 Int.	n	1954 Int.	n	Identification						
5071·0			0·5	1	?Cr IV	4*F*	71·6				
5084·0:	?1	1	1*N*	4	?Fe III	1*F*	84·8				
5147·9			1*n*?*d*	3							
5191·9	1	2	4	4	A III	3*F*	91·4				
5215·0			1*n*	2							
33·9	6	5	16	4				Fe II	49	34·6	°
5288·6	2	3	6	4							
5305·1:			1*n*	1							
5323·9			1·5	4	?Cl IV	3*F*	22·2				
5412·4	2*N*	3	5*W*	4	He II	2	11·5				
5494·1	1	2	1·5	4				Fe II	17*F*	95·8	*
5518·6	1	3	3	4	Cl III	1*F*	17·2				
35·2	1	2	1	1	Cl III	1*F*	37·7	Fe II	55	34·9	
44·2			?0·5	1							
52·4			?0·5	1	?Cr III	2*F*	50·3				
67·9			?0·5	1							
77·8	?1	1	2	3	?O I	3*F*	77·45				
5592·1			2	3	O III	5	92·4				
5666·6	?1*n*	1	0·5	2	N II	3	66·6				
5679·8	2*n*	2	2	3	N II	3	79·6	?Fe VI	1*F*	78·0	*
5783·6			1·5	3	?Cr III	1*F*	85·4	?Mn VI	2*F*	83·4	
5799·7	4*n*	3	7*W*	4	C IV	1	01·5				
5813·3:	1*n*	2	2*n*	2	C IV	1	12·1				
20·7	?1	1	3	4							
34·8			1	1							
53·9			?0·5	1							
5860·6			?0·5	1							
5904·5:			1	2	?Mn VI	2*F*	07·1				
5911·3	1	3	4	4							*
6047			1	1							
6101·9			5	4	K IV	1*F*	01·1				
6312·2	10	3	10	4	S III	3*F*	10·2				*
6731:	?1	1			?S II	2*F*	31·3				
6793·2	3	4	2:	2	K IV	1*F*	94·8				*
6997	?1	1	4	2	?Ti III	3*F*	91·8	?A V	1*F*	06·3	
7065	20	2	35	2	He I	10	65·2				
7101:	0·5	2	0·5	1							
36	10	1	15	2	A III	1*F*	35·8				
55			2	2	Fe II	14*F*	55·1	?Ti III	3*F*	52·8	
71	2	1	5	2	A IV	2*F*	69·0	He II		77·5	°
7189	?0	1	3*n*	2	Fe IV	*F*					
7207			?0·5	1							
21	?0	1	3	2	Fe IV	*F*					
38			3	2	A IV	2*F*	36·0				
62			2	2	A IV	2*F*	63·3				
7281	4	1	6	2	He I	45	81·3				
7324	8	1	10	2	O II	2*F*	(blend)	A IV	2*F*	32·0	°
7378			1*n*	2	C IV			Ni II	2*F*	77·9	*
7749	4	1	8	1	A III	1*F*	51·0				
7816	?0	1	?1	1	He I	69	16·2				
7923			1*n*	1							
8046			1*n*	1	Cl IV	1*F*	46·1				
8361	2	1	?1*n*	1	He I	68	61·8				
8446	10	1	10	1	?O I	4	46·5				

NOTES TO TABLE II

3490 On the Newtonian spectrogram (120 A/mm) this strong very diffuse feature is barely resolved from 3478 and the measured wave-length is very uncertain.

3721·8 The contribution of [S III] to this line is indicated by comparing its intensity with those of neighbouring Balmer lines.

3891·65 (Plate 2 (*a*)). Wave-length and intensity uncertain owing to proximity of H 3889.

3954·4 Identification very uncertain.

4097·3 Strengthened from 1953 to 1954 ; well resolved from Hδ (Plate 2 (*b*)). Not recorded in Paper II but, like N III 4640, was probably present, very diffuse, blended with the wing of Hδ.

4152·46 (Plate 2 (*b*)). Unidentified. Observed also in η Car. The ascription to [Co II] in Paper II is now ruled out by the subsequent behaviour in RR Tel.

4206·62, 4209·04 Unidentified. Observed in RR Pic as a single line 4207. Assuming the two lines to have a common origin the separation may provide a clue to the identification.

4571·24 Can scarcely be due to Mg I at this stage of ionization.

4868·2 (Plate 2 (*c*)). Wave-length and intensity uncertain owing to proximity of Hβ. Identification with [Fe IV] due to Edlén (private communication).

4930·8 Probably due to Fe III (1*F*). [O III] 4931·8 should be several thousand times fainter than the nebulium pair.

5233·9 Fe II must be a minor contributor. This unidentified line appeared in RR Pic.

5494·1 Mainly unidentified ; this line is too strong for Fe II 17 *F*, and the wave-length disagrees.

5679·8 The absence of 5177 shows that the contribution of [Fe VI] must be small.

5911 Observed in N Her.

6312·2 As in Paper II the wave-length disagrees from the prediction for [S III] but there can be little doubt of the correctness of the identification.

6793·2 The lower intensity recorded in 1954 is a result of the insensitivity of the emulsion used. In 1953 103aF emulsion recorded this line well, but not the other [K IV] line at 6101.

7171 In 1953 this line was regarded as too strong for [A IV]. In 1954 it appears stronger than the other [A IV] lines ; probably He II 7177 contributes.

7324 (Plate 2 (*d*)). Just resolved but measured as a single line.

7378 Probably due to C IV as observed in planetary nebulae (see Swings and Jose, *P.A.S.P.*, **61**, 181, 1949) with [Ni II] as a minor contributor.

3. *Representation of elements*

H The Balmer series is represented from Hα to H 18 (3691). A continuum appears on the Newtonian spectrogram extending well beyond the series limit. The Paschen series is recorded from 8863 to 8500 A. The stronger lines have very wide faint wings—see Hβ on Plate 2 (*c*).

He I New lines due to this element appear at 7816, 7281, 8361.

He II 4686 still appears as a very strong wide band and 5412 is also strongly winged.

C II 4267 is weakly present.

C III 4650 has been increasing in intensity since 1951.

C IV 5801, 5812, a well resolved pair despite strong wings, became quite a prominent feature in 1954. 5801 appeared weakly in 1952 but was regarded as unidentified in Paper II. 7378 appears weakly.

N II Weakly present ; multiplet 3.

[N II] Still prominent.

N III Multiplets 1, 2, 17. The 4640 group appeared as a very diffuse, unresolved band in 1952 (much stronger than in 1951); in 1953–54 all N III lines appeared much sharper, with 4634, 4640 and 4097, Hδ as well resolved pairs.

N IV The strong multiplet 1 is the chief feature in the far ultra-violet. The very broad band (4048 to 4064) recorded as probably due to N IV in 1952 has disappeared. This feature probably has the same origin as the broad N III bands.

O I The line 8446, apparently excited by Ly β (Bowen, *P.A.S.P.*, **59**, 196), appears strongly in 1953 and 1954. These infra-red plates are much better exposed than those of Paper II, and O I 8446 seems to have weakened relative to [O II] 7324, in agreement with the general trend of increasing ionization.

[O I] 6300, 6363 are still prominent. In addition, the auroral line 5577 appears to be weakly present.

[O II] The pair at 7324 is prominent, but the 3727 pair cannot be detected even on the Newtonian spectrogram. No permitted lines are observed.

O III Possibly represented by the one line 3757.

[O III] 5007, 4959 and 4363 are still very broad bands.

[F II], [F III], [F IV] Not found.

[Ne III] 3869, 3967 present as very strong bands, like [O III], (Plate 2 (*a*)). 3343 appears on the Newtonian spectrogram (1954).

[Ne IV] Multiplet 1*F*, with uncertain wave-lengths near 4720, is apparently absent but might be expected to appear soon.

Mg I The continued presence of the line 4571·2, attributed doubtfully to Mg I in Paper I, means that some other identification should now be sought.

Mg II 4481 still appears weakly in 1954.

Si II Represented by multiplets 1, 2, ?4, 5.

Si III, IV Not found, rather surprisingly in view of the presence of Si II.

[S II] Represented by 1*F* and ?2*F*. The weakness or absence of 6717, 6731 closely parallels the weakness or absence of [O II] 3727.

[S III] 6312 (3*F*) persists; 3797, 3721 (2*F*) are both blended with H lines, but appear to be present judging by the Balmer decrement.

[Cl III] 5517, 5537 (1*F*) were recorded in 1953 and 1954.

[Cl IV] A line measured four times in 1954 at 5323·9 may be due to multiplet 3*F* (5322·2). 8046 (1*F*) was measured once.

[A III] Represented by 1*F*, 3*F*. The strengthening of these lines is one of the outstanding features of recent developments. In Paper II 7135 (1*F*) was weakly recorded.

[A IV] Represented by 1*F*, 2*F*. 4740 was recorded weakly as an unidentified line in 1952.

[K IV] Represented by 6101, 6793 (1*F*) and 4511 (2*F*). Only 6793 was found in 1953.

Ca, Sc, V Not found.

[Ti III] Doubtfully represented by 6992, 7153 (3*F*), but the wave-lengths are discrepant.

[Cr III] Doubtfully represented by weak 5783 (1*F*) and 5552 (2*F*) with discrepant wave-lengths.

[Cr IV] Doubtfully represented in 1954 by a very weak line once measured at 5071 (4*F*).

[Mn IV] Possibly represented by 4529 (2*F*) which increased in intensity from 1953 to 1954. This line is analogous to 4658 of [Fe III]. The next strongest line is blended with Mg II 4481.

[Mn VI] Very doubtfully represented by two members of 2*F*, but 1*F* should appear too.

Fe II, [Fe II] Still represented by very numerous lines. The permitted lines have faded considerably. The progressive change in ratio between permitted and forbidden lines is well illustrated by the lines 4233, 4244 whose intensities for each year since 1949 are as follows:

		1949	1950	1951	1952	1953	1954
Fe II	4233	4	4	8	6	3	3
[Fe II]	4244	...	2	6	7	8	10

The absolute intensities are not of course strictly comparable from year to year; the intensities for 1949–50, dependent only on Newtonian spectra, should certainly be increased for comparison with later years.

Fe III A line at 3954 is very doubtfully ascribed to this ion, but the more familiar line 4419 does not appear.

[Fe III] Still represented by multiplets 1*F*, 3*F*.

[Fe IV] A special search for the multiplets in green and infra-red, predicted by Edlén* without precise wave-lengths, was conducted in 1953 and 1954. A number of unidentified lines, found in 1953, were listed as possibly due to this ion.† I am much indebted to Professor Edlén for communicating privately that the following lines in RR Tel can be definitely ascribed to [Fe IV] according to the independent analysis of his laboratory data:

Observed wave-length	Int. (1954)
4868·2	8:
4900·2	8
4906·2	12
4918·1	4
7189	3n
7221	3

* B. Edlén, Paris Colloquium on Novae and White Dwarfs, 1939.

† A. D. Thackeray, *The Observatory*, **74**, 90, 1954.

The lines 4868 and 4906 suffer from blending with Hβ and Fe II respectively (Plate 2 (*c*)). The other strong line 4900 has been observed in η Car (Paper II) and in other objects. RR Tel, with its sharp lines and continued brightness, is by far the best object in the sky for the astrophysical study of [Fe IV] lines.

[Fe V] Members of multiplets 1*F*, 3*F* appeared weakly in 1954. All lines observed are listed in Table II, (Plate 2 (*a*)).

[Fe VI], [Fe VII] Not found. N II 5679 masks [Fe VI] 5678.

[Co II] The line at 4152 (Plate 2 (*b*)), previously observed in η Car and RR Tel, and tentatively ascribed in Paper II to [Co II] on the ground of good agreement in wave-length and of relatively high transition probability in the region studied, must now be regarded as unidentified in both objects. The increase in strength in 1953 and 1954 points to an ion of higher ionization potential.

[Ni II] The few lines previously recorded have faded. In 1954, 4201 alone was once recorded weakly.

Ni III No coincidence with lines in RR Tel could be found with Velasco's unpublished laboratory lines due to this ion. The identification of some lines with [Ni III] or [Ni IV] however remains an interesting possibility.

The strongest unidentified lines are 4152, 4206, 4209, 5033, 5041, 5233, 5288.

4. *Summary of developments in* 1953–54.—The chief characteristic of the development of RR Tel since spectroscopic observations began in 1949 has been a gradual increase in the state of ionization. In 1953–54 this is exemplified well by the emergence of C IV, [Fe IV], [Fe V], [A III], [A IV], [Cl IV], [K IV]; there has been no very marked change in individual lines like the emergence of strong broad bands of [O III] and [Ne III] which occurred between the 1951 and 1952 seasons.

A second feature is the fading of Fe II permitted relative to forbidden lines (see preceding section). In this connection it is worth emphasizing the continued presence of Si II lines but non-appearance of Si III and Si IV. The permitted lines of higher stages of ionization which have recently appeared are confined to the light elements C, N, etc. The absence of Si II 4128, 4130 suggests dilution of radiation.

Thirdly, there is a striking difference in character between lines of low and high stage of ionization or excitation. Fe II lines are sharp, while lines such as [Fe V], [K IV] 4511 are quite diffuse. (See Plate 2 (*a*, *b*) for comparison of Fe II and [Fe V]). Among the stronger lines, the series of increasing diffuseness is continued by C IV 5801, [O III] and [Ne III] and the broadest band of all, He II 4686. Thus diffuseness appears to be correlated with both increasing ionization and excitation potential.

5. *Rate of development compared with other slow novae.*—It is well known that different novae tend to show roughly the same sequence of spectroscopic developments, but that the rate of development varies markedly. McLaughlin*, in particular, has derived quantitative relationships between the rates of development and of decline in light and the absorption velocities of novae.

* D. B. McLaughlin, *Ap. J.*, **85**, 362, 1937; **91**, 369, 1940; **95**, 428, 1942.

Sufficient spectroscopic observations of RR Tel during the nebular stage have now been accumulated to permit a fair comparison with the corresponding stages of the well-observed slow nova RR Pic.* The observations of spectroscopic changes in RT Ser and η Car are very scanty, but nevertheless an attempt to compare these other two slow novae with RR Tel is of some interest.

The comparison of stages of spectroscopic development can be supplemented by two points on the light-curves. In the cases of RR Pic, RR Tel and RT Ser it is known that the epoch of maximum light was preceded by a steep rise some months or years earlier; between these dates the light-curves are fairly flat at a magnitude or so fainter than the brightest maximum. The preliminary rise of RR Pic was not actually observed, but must have occurred between 1925 January 13 and April 13.† η Car was at its brightest known maximum in 1843 and is known to have fluctuated between 2nd and 4th magnitude from 1677 to 1830.

The main object of the accompanying Table III is to compare the various developments during the nebular stages. Consequently the zero-point of the time scale is chosen to coincide with the "Eta Carinae" stage ([Fe II] predominant), although in η Car itself this stage seems to have persisted without much change for at least 40 years. Many adopted stages correspond to marginal appearances of various states of ionization, especially of Fe. The dates of marginal appearance are of course subject to very great uncertainty, especially when comparing spectra taken with widely different spectrographs. It must be remembered too that the marginal appearance of any emission line will depend upon the strength of any continuous spectrum; however, the continuum is usually very weak in these objects at the corresponding phases. The marginal appearance of [Fe III] in η Car is given as 1951, since this ion is well represented in Table II of Paper II, while Gaviola's observations‡ from 1944 to 1951·2 do not include lines such as 4734, 4754 and there is no certain evidence of the ion at all in earlier years.

As was pointed out in Section 3, the development of RR Tel has involved the more rapid fading of Fe II than of [Fe II] lines. The ratio of permitted to forbidden line intensities (e.g. 4233/4244) seems to vary considerably from nova to nova at the same stage of ionization. No doubt this ratio is an indication of a more complex set of physical conditions than the criteria of development used in Table III.

The last column of Table III gives an average time-sequence, based on all four objects, for the development of the various stages, in arbitrary units.

In Fig. 1 the data of Table III are plotted with RR Pic used as a basic comparison. It will be seen that there is a roughly linear relationship between the rates of development of all four novae, surprisingly close considering the uncertainties of the dates. The slopes of the straight lines drawn through the points correspond to the following time-scales for corresponding stages of development of the four novae:

RR Pic	RR Tel	RT Ser	η Car
1·0	6	17	220

The above figures are necessarily rough, their probable errors being of the order of 25 per cent.

* H. Spencer Jones, *Ann. Cape Obs.*, X, pt. 9, 1931.
† H. Shapley, *H.B.*, 823, 1925.
‡ E. Gaviola, *Ap. J.*, **118**, 234, 1953.

TABLE III

Development of four slow novae

Stage	RR Pic	Δt (days)	RR Tel	Δt (yrs)	RT Ser†	Δt (yrs)	η Car	Δt (yrs)	Δt (Arbitrary units)
Preliminary rise	1925 Jan. 13 1925 Apr. 13	−284 −194	1944·9	−5·8	1909	−19			−7
Maximum	1925 June 10	−136	1948·5 :	−2·2 :	1920	− 8	1843	−77 :	−3
cF absorption (Ti II !)	1925 Sept. 10	− 44	1949·4	−1·3			1893‡	−27 :	−1
η Car	1925 Oct. 24	0	1950·7	0	1928	0	1900 −1940	0 :	0
[Fe III] m.a.	1925 Dec. 1 :	+ 38 :	1951·5	0·8	1931	+ 3	1951	+ 31 :	+ 1
[O III] m.a.	1926 Feb. 1 :	+100 :	1951·5	+0·8					+1·5
He II flash	1926 Mar. 15	+142	1952·4	1·7					+2·5
[Fe IV] m.a.	1926 Apr. 28	+186	1953·3	+2·6					+3·5
[Fe V] m.a.			1954·3	+3·6					
[Fe VI] m.a.	1926 June 4	*+223			1940	+12			+5
[Fe VII] m.a.	1926 Sept. 7	*+318			1942	+14			+6

* From Spencer Jones (*op. cit.*, Table XLIV, page 134 1). The marginal appearances are taken from [Fe VI] 5146, [Fe VII] 5720. The stronger lines 5177, 6085 are liable to blending with Fe II and [Ca V] respectively at earlier dates.

† Data for RT Ser are taken from S. & C. P. Gaposchkin, *Variable Stars*, pp. 263, 264, and Struve and Swings, *Ap. J.*, **92**, 295, 1940; **96**, 471, 1942.

‡ B. J. Bok (*P.A.*, **38**, 399, 1930). In 1893 η Car was nearly a magnitude brighter than in preceding years, and it is not entirely certain that a bright-line spectrum may not have been prominent earlier. In 1886, however, the spectrum was observed visually to be continuous plus absorption lines (Clerke, *The Observatory*, **11**, 429, 1888).

It is of interest to note that since McLaughlin found that N. Lac 1936 developed at 16 times the rate of RR Pic, this rate corresponds to about 3500 times that of η Car. However, no close parallelism can be expected between the developments of the fastest and the slowest novae; it is the slow novae like RR Tel that seem to exhibit the purest examples of sequences of increasing ionization.

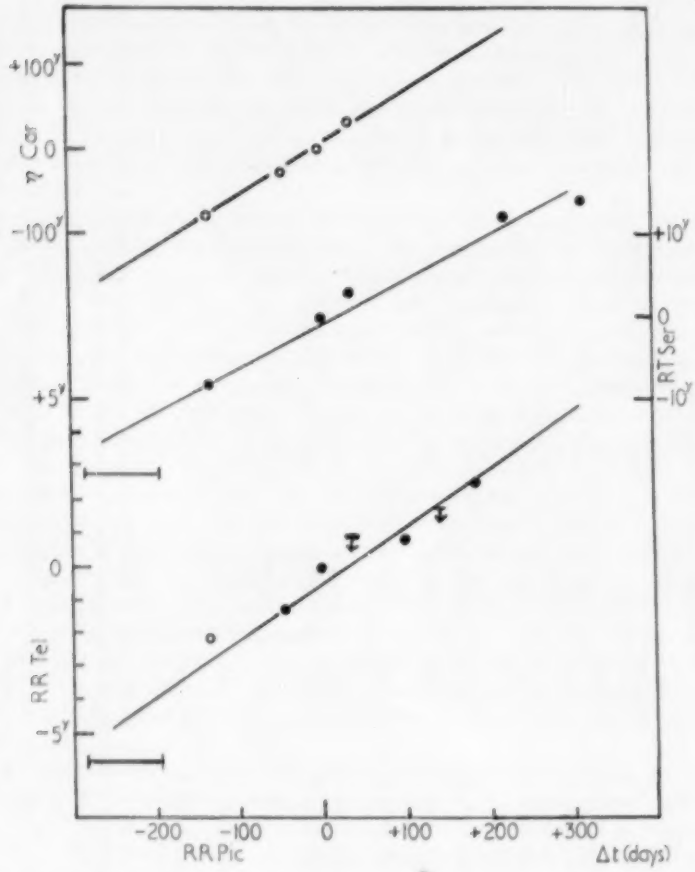


FIG. 1.—*Epochs of corresponding stages of development of RR Pic (abscissae) and η Car (top), RT Ser (middle), RR Tel (bottom).*

The close relationships exhibited in Fig. 1 suggest that extrapolations of the lines may be attempted without falling into the realm of pure speculation. Thus, by extrapolating the line for RR Tel we find that the marginal appearance of [Fe VI] and [Fe VII] may be expected at about 1954·8 and 1955·8 respectively. No sign of [Fe VI] 5177 could be detected on the last plate of that region (1954·76) but its appearance in the near future followed by [Fe VII] may be confidently expected.

Extrapolation of the line for η Car suggests that the [O III] lines should appear about 1980. As was pointed out in Paper II, however, the lines of oxygen in η Car seem to be surprisingly weak. It will therefore be of special interest to see whether the [O III] lines do follow the pattern of development suggested by other novae.

The first two entries in Table III refer to points on the light-curves and their relation to the other entries is probably insecure. Nevertheless it is of interest to attempt to estimate from them when a preliminary rise of η Carinae may have occurred. A straightforward application of the relative time-scales suggests a date between 1670 and 1710, with an uncertainty of about half a century. Halley's observation in 1677, when the star was ranked as 4th magnitude, is usually accepted as the first record, and one might therefore suppose that a preliminary rise from pre-nova state occurred shortly before. However, Bayer's *Uranometria* (1603) apparently recorded the star as 2nd magnitude on second-hand evidence.* As pointed out by Innes†, the star was probably fainter than 4th magnitude in Ptolemy's time, for it was not included in the *Almagest*. Any records of the star prior to 1677 would be valuable in this connection and the problem seems to deserve the attention of historians of astronomy.

* Quoted by Agnes Clerke, *Problems in Astrophysics*, p. 371, 1903.

† *Ann. Cape Obs.*, IX (not XI as sometimes quoted), 75B.

Radcliffe Observatory,
Pretoria :
1955 *February* 2.

ADDENDUM

Note added in proof. I am deeply indebted to Dr A. Armitage who has supplied the following information concerning the history of Eta Carinae, based on Muller and Hartwig, *Geschichte u. Lit. des Lichtwechsels*, Leipzig, 1918, 1920.

The non-inclusion of Eta Carinae in the oldest Catalogues (Ptolemy, etc.) allows no conclusions to be drawn in view of the fact that bright stars such as μ Argus (10° farther north) are not included. The earliest records of the region of Eta, depending on observations in India at the end of the sixteenth century, are included in Bayer's and v. Houtmanns' charts. v. Houtmanns' (1603) does not show the star, while he includes p, q, s, μ Car, whence Eta can hardly have been brighter than 5^m. Bayer (1603) includes a group of some seven stars, designated d, of 4th magnitude; the modern Eta may have belonged to this group, but Bayer's Eta Argus, of 2nd magnitude, is certainly not Eta Car.

Lacaille (1757) assigns 2nd magnitude to his Eta Argus, with an accurate position which clearly identifies it with the modern Eta Car.

One may conclude that the star must have brightened at some time between 1600 and 1757, but that the richness of the Milky Way region renders naked-eye records practically unusable for estimating the extent of the brightening.

EMISSION OBJECTS IN THE SMALL MAGELLANIC CLOUD SHOWING THE N_1, N_2 NEBULAR LINES

E. M. Lindsay

(Received 1955 March 31)

Summary

Out of a total of over 100 objects in the Small Magellanic Cloud discovered on ADH spectrum plates showing Hα emission, thirty-three have the N_1, N_2 nebular lines in their spectra. These latter have been divided into two groups. The members of Group I, containing thirteen objects, are associated with nebulosity and are doubtless early-type stars exciting the surrounding nebulosity. Their mean absolute photographic magnitude is $-4{\cdot}9 \pm 0{\cdot}6$ (p.e.) and mean colour index (blue *minus* red photographic) $-0{\cdot}47 \pm 0{\cdot}19$. The members of Group II, containing seventeen objects, are not associated with nebulosity and have a mean absolute photographic magnitude of $-3{\cdot}7 \pm 0{\cdot}6$ and mean colour index $+0{\cdot}77 \pm 0{\cdot}29$. Three objects were too bright to determine magnitudes on the plates available. It is suggested that the members of Group II are planetary nebulae. An additional eleven objects showing Hα emission and with the same characteristics of Group II are possible planetaries.

A list of emission objects published by Miss Cannon (**1**) in 1933 contains only thirteen such objects as being then known to be present in the Small Magellanic Cloud. Of this number, eleven were classified as nebulous and two as O-type stars.

Using a 10·5-in. visually corrected camera and 15° objective prism of the Mt Wilson and Palomar Observatories, the Lamont–Hussey Observatory of the University of Michigan have completed a survey of the Small Magellanic Cloud in Hz light. By this method 120 new emission objects have been discovered by Henize and Miller (**2**). Eighty-eight of these are classified as "stellar" and thirty-two as "nebulous". As pointed out by the authors, criteria were lacking for further classification of the emission objects classed as stellar, so that although the nebulous objects would include only gaseous nebulae the stellar objects may include small nebulous knots, planetary nebulae and a variety of stellar types.

A catalogue of nebulosities in the Small Cloud has been compiled by Virginia Nail, Whitney and Wade (**3**). The instrument used was the 3 in. *f*/1·5 Zeiss–Sonnar camera of the Boyden Station stopped to *f*/2, with a red glass filter (Corning 2403) and Hα interference filter. There were also available for examination a 90^m red (103a–E) plate, a 60^m blue (103a–O) and a red (103a–E) objective prism plate, all taken with the ADH telescope. The nebulosities were classified according to the system of Bok and Wade (**4**). In addition to their types, a Type V was added which indicates objects with stellar images. This type contains 41 per cent of the total of 152 catalogued nebulosities.

In tests carried out by the writer on the ADH prism (angle 3° 10′; dispersion at Hγ 240 A/mm) attention was concentrated on two regions, the Large and the Small Magellanic Clouds. The large amount of material made available by the tests enabled a survey to be made of bright-line objects in the Clouds. The survey is completed for the Small Cloud; additional plates are needed to complete the survey for the Large Cloud. Since it was known that a search for nebulosities in these regions was being made at Harvard, no attempt has been made to catalogue diffuse nebulosities or stellar but nebulous looking objects. So as to avoid overlapping of the work at Harvard and Armagh the survey has been confined, with a few exceptions, to stellar-looking objects showing emission lines in their spectra.

A total of 106 emission objects has been catalogued for the Small Cloud. The present paper deals only with those which show the N_1, N_2 nebular emission lines (λλ 5007, 4959). The observational material is as follows.

Spectra.—The most fruitful discovery plates were red plates (103a–E) used in conjunction with a Wratten No. 25 A red filter. The filter cuts off sharply all light to the red of Hα. Exposures varied from 30^m to 2^h, but nearly all the emission-line objects are registered on a 60^m exposure plate of excellent quality. In addition to these, red plates (103a–E) without filter and blue plates (103a–O and IIa–O) were available with exposures up to 90^m, including widened and unwidened spectra. Guiding was carried out by means of a "gooseneck" micrometer eyepiece fitted to the 8-in. guiding telescope attached to the main tube of the ADH. Widening was obtained by guiding on the cross-wire of the eyepiece as for direct photography, periodically offsetting the cross-wire by a known number of turns of the micrometer screw and immediately bringing back the guiding star to the cross-wire by the slow-motion guiding mechanism.

Magnitudes.—Blue photographic magnitudes have been determined from eight ADH IIa–O plates with exposures ranging from 5^s to 1^h; red magnitudes from six ADH 103a–E plates with the 25 A red filter, the exposures ranging from 1^m to 90^m. These plates were taken on a programme for Dr Shapley and I am grateful to him for permission to use them first at Armagh. The magnitudes were estimated by means of a scale of stellar images taken in the blue and red with the ADH. The scale was calibrated for each plate by means of King's unpublished magnitudes of some of the stars in Sequence I of the Small Cloud (**5**). King's magnitudes were determined with the photoelectric photometer attached to the 60-in. Rockefeller reflector, and are reduced to the International Scale. His faintest magnitudes in the blue and red are $m_{pg} = 15{\cdot}06$ and $m_{pr} = 15{\cdot}64$ respectively, which are approximately the limiting magnitudes on a 30^s and 5^m exposure in the blue and red respectively with the ADH. The mean difference between King's magnitudes and those given in *H.A.*, **108** for the four faintest stars (t, v, w and x) in King's list is (*H.A.*—King) = $+0^m{\cdot}20$. Stars fainter than those in *H.A.*, **108** were corrected by $-0^m{\cdot}2$ and used for the long-exposure blue plates. King's red magnitudes are the only ones available in the Small Cloud. It was therefore necessary sometimes to extrapolate the calibration curves slightly to reach beyond his limit. In some cases small corrections, not exceeding $0^m{\cdot}2$, were applied to the measured magnitudes for distance from the standard stars in accordance with the preliminary estimates for distance correction found by the writer (**6**).

The objects showing nebular lines are listed in Table I, which gives the catalogue numbers, right ascensions and declinations (1900), the positions in X and Y coordinates as described in *Harvard Annals*, **60,** No. 4 (1907), and spectral lines. The right ascensions and declinations are rough measurements derived from graph paper with the use of a few standard stars.

The objects in Table I have been re-arranged in Table II which gives the blue and red photographic magnitudes, the colour indices (blue *minus* red photographic), the photographic absolute magnitudes and remarks. The modulus $(m - M)$ of the Cloud has been taken to be 18·9 photographic (**7**). No correction has been applied to allow for space absorption.

TABLE I

Objects in the Small Magellanic Cloud with nebular lines in their spectra

No.	R.A.		Dec.		X	Y	Emission lines in addition to Hα, N_1, N_2
	h	m	°	′			
1	0	28·6	−72	16·7	6552	13260	Hβ, Hγ
2	0	37·5	73	21·9	9264	9642	Hβ
3	0	37·7	74	22·1	9522	5916	Hβ
4	0	39·6	73	11·2	9792	10092	
5	0	39·9	73	35·3	9918	8652	Hβ, Hγ, Hδ, Hε, Hζ, λ 3727
6	0	41·3	73	49·7	10338	7800	Hβ, Hγ
7	0	41·8	73	56·0	10440	8520	Hβ, Hγ
8	0	41·8	73	37·7	10464	7428	Hβ, Hγ, Hδ, λ 3727
9	0	42·9	73	39·0	10704	8454	Hβ
10	0	44·7	73	51·1	11142	7944	Hβ, Hγ
11	0	45·9	73	40·8	11490	7242	Hβ, λ 4363, Hγ
12	0	46·2	74	17·7	11640	6180	Hβ, Hγ
13	0	47·6	74	31·3	11976	5388	Hβ, Hγ, Hδ
14	0	48·1	73	46·2	12072	8040	Hβ, Hγ
15	0	48·2	73	24·0	12072	9486	Hβ, Hγ, Hδ, λ 3727
16	0	48·3	71	57·1	12012	14628	Hβ, Hγ
17	0	54·1	73	05·4	13596	10548	Hβ, Hγ, Hδ, λ 3727
18	0	54·9	73	11·8	13806	10158	Hβ, Hγ, Hδ, λ 3727
19	0	55·1	73	12·7	13848	10098	Hβ, Hγ, Hδ, λ 3727
20	0	55·4	73	00·5	13926	10848	Hβ
21	0	55·7	72	42·8	14106	11838	Hβ
22	0	55·9	72	43·1	14070	11910	Hβ
23	0	57·6	72	08·0	14610	13980	Hβ, Hγ, Hδ, Hε, λ 3727
24	0	59·5	72	25·8	15108	12894	Hβ, Hγ, Hδ, λ 3727
25	1	06·2	73	44	16650	8094	Hβ, λ 4363, Hγ, Hδ, Hε, Hζ, λ 3727
26	1	11·1	73	50·0	17832	7620	Hβ, Hγ, Hδ, Hε, Hζ, λ 3727
27	1	11·3	73	47·9	17892	7674	Hβ, Hγ, Hδ, λ 3727
28	1	11·4	73	47·5	17958	7752	Hβ, Hγ, Hδ
29	1	11·9	73	50·2	18036	7578	Hβ, λ 4686, Hγ, Hδ, λ 3727
30	1	12·0	73	51·2	18054	7578	Hβ, λ 3727
31	1	12·1	73	51·6	18072	7494	Hβ, Hγ, Hδ, λ 3727
32	1	18·4	73	46·4	19740	7650	Hβ, Hγ, Hδ
33	1	21·6	73	40	20538	7896	Hβ, λ 4686, λ 4363, Hγ, Hδ, Hε, Hζ, λ 3727

Table II includes two types of objects differing chiefly in their colour indices but also in their absolute magnitudes. In Group I the colour indices are negative; in Group II positive. Table III gives the mean colour indices and mean absolute magnitudes of each group together with their probable errors and the numbers in each group.

TABLE II

Magnitudes, red indices and absolute photographic magnitudes of the objects of Table I

No.	m_{pg}	m_{pr}	C.I.	M_{pg}	Remarks
7	14·02	14·57	−0·55	−4·9	Hα large and diffuse. Region of extended nebulosity
8	14·23	14·41	−0·18	−4·7	Hα strong and diffuse. Star in a concentration of an oval-shaped nebulosity
9	13·25	14·10	−0·85	−5·6	Hα large and diffuse. Star involved in nebulosity which is connected with the nebulosity of No. 8
14	14·23	14·67	−0·44	−4·7	Hα diffuse and oval. Small cluster probably with nebulosity
18	14·64	15·41	−0·77	−4·3	Circular nebulosity with central star. An H II sphere. Harvard No. 97
21	15·35	15·41	−0·06	−3·6	Very nebulous region
22	13·98	14·13	−0·15	−4·9	Very nebulous region
24	14·84	15·33	−0·50	−4·1	Nearly symmetrical circular nebulosity with central star. An H II sphere. Hα is double. On long-exposure red plates there is adjacent faint nebulosity. This is No. 188 in the Shapley–Wilson Catalogue (8) where it is described as " nebulous star, planetary ? "
27	14·17	15·03	−0·86	−4·7	Very nebulous region
28	15·16	15·50	−0·34	−3·7	Circular nebulosity with central star. An H II sphere. Harvard No. 149
29	11·60	12·40	−0·80	−7·3	Very nebulous region
30	13·55	14·00	−0·45	−5·3	Very nebulous region
31	12·75	12·87	−0·12	−6·2	Very nebulous region
1	15·92	15·67	+0·25	−3·0	Hα very faint. No continuous spectrum
2	16·19	15·74	+0·45	−2·7	Hα very faint. No continuous spectrum
3	16·29	15·48	+0·81	−2·6	Hα very faint. No continuous spectrum
4	16·21	15·79	+0·42	−2·7	Hα very faint. No continuous spectrum
5	13·75	13·11	+0·64	−5·2	Hα moderately strong. No continuous spectrum
6	14·40	13·93	+0·47	−4·5	Hα broad and diffuse, possibly double. No continuous spectrum. Overlapping a fainter star. Very nebulous on red plates. Harvard No. 16, described as a " bright nebulous star " of Type V
10	15·15	13·79	+1·36	−3·7	Hα strong. No continuous spectrum. Although situated between an H II sphere and extended nebulosity it is apparently not associated with any nebulosity
11	14·60	13·07	+1·53	−4·3	Hα moderately strong. No continuous spectrum
12	15·93	15·03	+0·90	−3·0	Hα moderately weak. No continuous spectrum
13	15·66	14·74	+0·92	−3·2	Hα moderately strong. No continuous spectrum
15	14·61	14·26	+0·35	−4·3	Hα moderately strong. Star nebulous. Faint continuous spectrum
16	15·59	15·16	+0·43	−3·3	Hα moderately faint. No continuous spectrum
17	14·23	13·27	+0·96	−4·7	Hα moderately strong. No continuous spectrum
19	14·64	14·01	+0·63	−4·3	Nebulous object in neighbourhood of five stars. There is an impression of faint nebulosity but 60^m exposures on blue and red plates taken with the 60-in. reflector show that at least part of this is due to fainter stars. Harvard No. 99 ($X=13839$; $Y=10081$) is described as " nebulous group, three nebulous stars ". Hα fairly strong and no continuous spectrum
20	15·65	14·37	+1·28	−3·2	Hα moderately strong. No continuous spectrum

TABLE II (*cont.*)

No.	m_{pg}	m_{pr}	C.I.	M_{pg}	Remarks
23	15·08	13·77	+1·31	−3·7	Nebulous object with possibly traces of nebulosity on long-exposure red plates. Not separated from a fainter star. Hα moderately strong with no continuous spectrum. Harvard No. 107, described as " nebulous star " of Type V. No. 182 in the Shapley–Wilson Catalogue and described as " nebulous star, planetary ? "
32	15·80	15·38	+0·42	−3·1	Hα moderately strong. Possibly faint continuous spectrum

Notes on Table II. No. 25 (IC 1644) is too bright to measure on the plates available (exposures 30^m blue and 90^m red). It is classified in *H.A.*, 60 as a gaseous nebula. There is no extended nebulosity in the immediate neighbourhood and it is difficult to say whether or not some adjacent faint stars give the impression of faint nebulosity to the object.

The emission from No. 26 comes from a condensation of stars and nebulosity. No measurements of magnitudes could be made.

No. 33 was off the field of any of the blue plates available. The value $m_{pr}=11·95$ was obtained from a single plate, and can only be approximate. It is described in *H.C.* 224 as a gaseous nebula. On a 90^m exposure ADH red plate with red filter it is a stellar object in a small cluster of much fainter stars. There is no noticeable nebulosity around it or in the neighbourhood. One might suspect a trace of faint nebulosity but, if so, it could as readily be interpreted as faint unresolved stars.

TABLE III

Mean values of Table II entries

Group	C.I.	*M*	No.	Remarks
I	−0·47 ± 0·19	−4·9 ± 0·6	13	Associated with nebulosity
II	+0·77 ± 0·29	−3·6 ± 0·6	17	Not associated with nebulosity

It seems clear that in Group I we are dealing with O- or early B-type stars which are exciting the surrounding nebulosity and giving the typical emission spectra of gaseous nebulae. It is impossible to classify the types since all are too faint or are in dense regions. Of the thirteen objects, two (Nos. 21 and 22) are in NGC 346 and five (Nos. 27–31) are in the neighbourhood of NGC 456 and NGC 465, confused regions with overlapping spectra and nebulosity. Three are circular nebulosities with a central star—so-called H II spheres. Data on these latter are given in Table IV and compared with the measurements of

TABLE IV

H II spheres in the Small Magellanic Cloud

(A = Armagh; H = Harvard)

Cat. No.		m_{pg}		C.I.		*M*	
A	H	A	H	A	H	A	H
...	30	13·91	13·81	−0·55	−0·65	−5·0	−5·1
...	75	15·00	14·90	−0·57	+0·34	−3·9	−4·0
...	79	...	15·60	...	−0·19	...	−3·3
18	97	14·64	14·93	−0·77	−0·07	−4·3	−4·0
24	115	14·83	15·25	−0·50	−0·33	−4·1	−3·6
...	119	15·7	15·42	−0·5	+0·94	−3·2	−3·5
28	149	15·16	14·95	−0·34	+0·15	−3·7	−4·0
Means (omit 79)		14·87	14·88	−0·54	+0·06	−4·0	−4·0

Virginia Nail, Whitney and Wade referred to above. The latter list seven such objects. The remaining four are not contained in Table I since they show only Hα emission. No measurements have been made on Harvard No. 79 since it appears on the Armagh plates as a nebulous object but with no trace of a central star. Harvard No. 75 was classified as probably a stellar object in a small cluster of faint unresolved stars but possibly nebulosity.

The Armagh and Harvard blue photographic magnitudes are in fair agreement. In four cases, however, the colour indices differ radically. Possibly part of this represents the difficulty of estimating magnitudes of stars involved in nebulosity. The divergences between the estimates by different observers of the magnitude of a comet with a strong coma are well known, and a good example of the difficulty is seen in the Armagh estimates of (for instance) the central star of the first object in Table IV. From exposure times of 20^s, 60^s, 3^m, 10^m, 30^m and 60^m the estimated magnitudes were 13·55, 14·05, 13·87, 14·18, 14·95 and 15·20 respectively. In the first four cases little or no nebulosity was evident and the mean of these was taken. On the last two exposures the nebulosity was strong and very strong and it was extremely difficult with this subjective method to estimate the magnitude of the embedded star. It was found in all cases that when the nebulosity became sufficiently dense to be troublesome, the estimates of the magnitudes became systematically fainter with increasing density.

Referring to Table I, it will be seen that three of the objects of Table IV (18, 24 and 28) show the N_1, N_2 lines and two also show λ 3727. It seems unlikely that these lines would appear with a central star of positive colour index. It is unexpected that the brightest star absolutely (Harvard No. 30; $M = -5{\cdot}0$) gives only an Hα sphere with no other lines present.

In Group II we have objects not associated with nebulosity, generally stellar in appearance, with little or no continuous spectrum, showing the typical discrete nebular lines and having positive colour indices. It is suggested that these are planetary nebulae. The values of the absolute magnitudes are higher than those found for planetaries in our own galaxy. Using data on mean parallaxes based on proper motions and the rotational terms of galactic rotation, Berman (**9**) and Parenago (**10**) have arrived at values of $-1{\cdot}5$ and $+0{\cdot}2$ respectively for the mean absolute magnitudes. The higher value of $M_{pg} = -3{\cdot}7$ found here may be due in part to the uncertainty in applying corrections for absorption of light in our galaxy and in part to a selection effect in the Small Cloud, where only the brightest planetaries would be found in the present material. With $M = -1{\cdot}5$ at the distance of the Cloud we would have $m = 17{\cdot}4$, which is fainter than any of the emission objects so far found.

The large positive colour indices are in keeping with the results of Vorontsov-Velyaminov (**11**) who has found a visual *minus* photographic index of $+1{\cdot}4$ for planetary nebulae from integrated magnitudes. This is explained by the composition of spectra of planetaries being mainly discrete emission lines and by the strength of the green nebular lines N_1, N_2. The strength of Hα will also explain the large positive colour indices derived from the difference between photographic and red magnitudes.

A number of objects have been found which have the same characteristics as the planetary nebulae above, namely Hα emission, little or no continuous spectrum, positive colour indices and comparable absolute magnitudes. These are listed in Table V with additional details in Table VI.

None of the objects in Tables V and VI is associated with any observable nebulosity. Two (Nos. 34 and 36) or possibly three (No. 44) are in or in line with small clusters. Taking all together, however, the mean colour indices and mean absolute magnitudes are $+0{\cdot}80 \pm 0{\cdot}12$ (p.e.) and $-3{\cdot}3 \pm 0{\cdot}5$ (p.e.) respectively. The mean colour index is naturally close to the value above,

TABLE V

Objects in the Small Magellanic Cloud which are possibly planetary nebulae

No.	R.A. h	R.A. m	Dec. °	Dec. ′	X	Y	Emission lines
34	0	41·6	−73	43·2	10284	8190	Hα faint, diffuse. No continuous spectrum
35	0	44·9	73	31·8	11244	8946	Hα faint. Suspected N_1, N_2. Hβ at plate limit. No continuum
36	0	45·4	74	04·0	11310	7446	Hα faint, diffuse. No continuum
37	0	46·2	73	42·8	11568	7584	Hα moderately weak and diffuse. Faint continuum
38	0	47·8	72	58·9	11946	10908	Hα weak. Suspected N_1, N_2. Hβ at plate limit. Possibly faint continuum
39	0	48·7	73	54·1	12138	7614	Hα faint. Very faint continuum
40	0	50·6	73	14·8	12696	9978	Hα strong. Hβ, λ 4686, Hγ. Generally λ 4686 is outstanding, but on one plate its intensity is equalled by that of Hβ with Hγ also strong. No continuous spectrum.
41	0	50·7	73	02·2	12726	10740	Hα weak and diffuse. No continuum
42	0	51·9	72	36·6	12990	12306	Hα fairly strong. Suspected Hβ, Hγ, Hδ at plate limit. No continuum ?
43	0	55·9	72	35·4	14094	11772	Hα faint. No continuum
44	1	12·9	73	41·5	18354	8082	Hα moderately strong and diffuse. No continuum ?

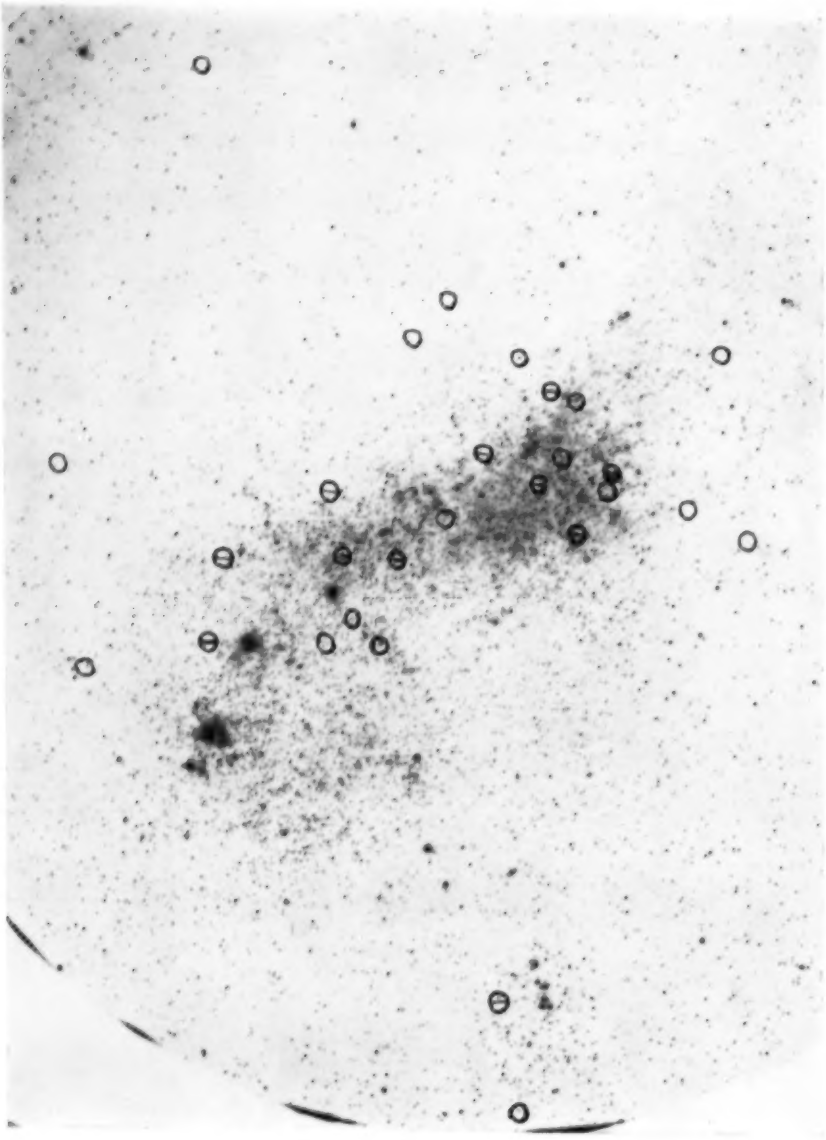
TABLE VI

Magnitudes, red indices and absolute photographic magnitudes of the objects of Table V

No.	m_{pg}	m_{pr}	C.I.	M_{pg}	Remarks
34	15·37	14·59	+0·78	−3·5	In a small cluster. No nebulosity on 90^m red or 60^m blue plate. Harvard No. 15 described as " nebulosity around four stars " of Type II
35	16·29	15·35	+0·94	−2·6	Somewhat nebulous faint star
36	15·46	15·09	+0·37	−3·4	In a small cluster. Some doubt of this object
37	14·96	14·34	+0·62	−3·9	Object nebulous
38	16·11	15·31	+0·80	−2·7	
39	16·10	15·22	+0·88	−2·8	
40	13·98	12·86	+1·12	−4·9	Near to but not to be confused with Harvard No. 79 which shows no emission lines
41	14·73	14·51	+0·22	−4·2	
42	15·72	14·09	+1·63	−3·2	
43	16·03	15·31	+0·72	−2·9	
44	16·23	15·48	+0·75	−2·7	Five stars close together. Difficult to be sure which one gives emission

since the objects were chosen partly for their positive indices; the mean absolute magnitude is somewhat lower, which is to be expected if they are too faint to show emission lines other than Hα. Taking all such objects in Tables II and V together we get C.I. $= +0{\cdot}78 \pm 0{\cdot}25$ and $M_{pg} = -3{\cdot}5 \pm 0{\cdot}5$. Their distribution with respect to absolute magnitude is given in Table VII. The sharp drop in the

W



N S

E

Distribution of probable (○) and possible (⊙) planetary nebulae in the Small Magellanic Cloud.

E. M. Lindsay, Emission objects in the Small Magellanic Cloud showing the N_1, N_2 nebular lines

TABLE VII

Distribution of planetary nebulae with respect to absolute magnitude

Limits of $-M$	5·5–5·1	5·0–4·6	4·5–4·1	4·0–3·6	3·5–3·1	3·0–2·6	2·5–2·1
No.	1	2	5	3	7	10	1

numbers at the faintest absolute magnitudes is doubtless due to working at the limit of the plates. Longer-exposure guided spectrum plates with the ADH will most probably reveal fainter objects. No. 33 has not been included in Table VII. If it were a planetary nebula with the average colour index of +0·8 its absolute photographic magnitude would be −6·0.

The positions of the objects suspected to be planetaries are plotted on Plate 3. No. 33 is just off the eastern edge of the plate. If it is included, the longest diameter is about 5°; the diameter at right angles to this is half that. There was not sufficient material for a complete search over the whole area of the Cloud as given by Shapley (**12**) from microphotometer tracings. One spectrum plate in the direction of the Large Cloud revealed no objects which could be classified as planetaries.

Armagh Observatory,
Armagh, N. Ireland:
1955 *March* 29.

References

(1) A. J. Cannon, *Harv. Bull.* No. 891, 1933.
(2) Karl H. Henize and Freeman D. Miller, *Publ. Obs. Univ. Michigan*, Vol. X, 1951.
(3) Virginia McKibben Nail, Charles A. Whitney and Campbell M. Wade, *Proc. Nat. Acad. Sci.*, **39**, No. 12, 1953. *Harv. Repr.* No. 377.
(4) B. J. Bok and Campbell M. Wade, Report of the Second Conference on the Gas Dynamics of Cosmic Clouds, Cambridge, England, 1953 (in press).
(5) Virginia McKibben Nail, *Harv. Ann.*, **108**, 1940. Ivan King, unpublished.
(6) E. M. Lindsay, *Irish Astr. Journal*, **2**, 140, 1953; *Armagh Obs. Leaflet* No. 17.
(7) Harlow Shapley and Virginia McKibben Nail, *Proc. Nat. Acad. Sci.*, **39**, 1953.
(8) Harlow Shapley and Harvia H. Wilson, *Harv. Circ.* No. 276, 1925.
(9) L. Berman, *Lick Obs. Bull.*, **18**, 57, 1937.
(10) P. P. Parenago, *Astr. J. Soviet Union*, **23**, 69, 1946.
(11) B. A. Vorontsov-Velyaminov, "Gaseous Nebulae and Novae", *Acad. Sc. U.S.S.R.*, Moscow and Leningrad, 1948.
(12) Harlow Shapley, *M.N.*, **94**, 791, 1934.

INTERPRETATION OF FRAUNHOFER-LINE PROFILES

H. H. Plaskett

(Received 1955 March 23)

Summary

From the experimentally determined cross-sections of the Na atom for collisions of the second kind it is shown that the hypothesis of true absorption is probably not applicable to the formation of Fraunhofer lines in the solar photosphere. Using atomic transition probabilities for radiative excitation and ionization, values of the source functions and the ratio of the line to the continuous absorption coefficients are derived in terms of kinetic, excitation and ionization temperatures. The functional dependence of these quantities on the optical depth is such as to permit a numerical integration of the exact equation of transfer, and the method of effecting this integration is outlined.

Introduction.—A satisfactory interpretation of Fraunhofer-line profiles, involving the minimum of hypothesis, requires a numerical integration of the equation of transfer for line formation. Unfortunately the form of the equation itself has not yet been settled. Thus in recent years, following Unsöld's adoption (**19**) of the heuristic and simplifying assumption of true absorption for the formation of lines in the spectrum of τ Scorpii and the justification he offered for its application to the solar photosphere (**20**), some workers have applied the resulting simple equation of transfer to find new values for the temperature distribution in the upper photosphere from the limb darkening of Fraunhofer lines. K.-H. Böhm (**2**) has, however, recently shown that, even allowing for possible thermal inhomogeneities due to granulation and for departures from local thermodynamic equilibrium, the thus derived temperature distributions fail to account for the limb-darkening both of faint lines and of the wings of the Na *D*-lines.

Invaluable though Unsöld's hypothesis has proved and is likely to prove in the interpretation of stellar-line profiles, it is probably not strictly applicable to the solar photosphere. It is the purpose of this paper to propose a method, a refinement of one earlier given (**13**), whereby a numerical integration of the exact equation of transfer can be effected. The paper is divided into three sections, the first of which examines the hypothesis of true absorption. In the second section values of the absorption and emission coefficients in the exact equation of transfer are derived, and in the final section these values are shown to lead to a simple method of numerical integration of the equation.

The essential part of the paper is contained in the second section. Upon its results, which must be demonstrably correct, depends the method of integration. It is suggested, however, that the simple ideas involved may be most readily grasped if the paper is read in the order Section 1, Section 2.1 and Section 3; reference can then be made back as necessary to the remaining sub-sections of Section 2.

1. *Hypothesis of true absorption*

1.1. *Test with the Na atom.*—In the strict sense of the words true absorption means that the line quantum absorbed is converted by a collision of the second kind then and there into kinetic energy of the free electrons and atoms at that place in the photosphere. If $\sigma_2(\epsilon)$ be the cross-sectional diameter for a collision of the second kind, then the number of such collisions experienced per sec by an atom of mass m_1 is given by

$$\frac{2n_2}{(kT)^2}\sqrt{2\pi kT\left(\frac{1}{m_1}+\frac{1}{m_2}\right)}\int_0^\infty \sigma_2^2(\epsilon)\epsilon\exp(-\epsilon/kT)\,d\epsilon, \qquad (1.1)$$

where n_2 is the number per cm^3 of colliding particles, each of mass m_2, and ϵ is the relative kinetic energy of the particles. From the experiments of Christoph (**4**), see Bates, Fundaminsky, Leech and Massey (**1**), the values of $\sigma_1(\epsilon+\epsilon_{12})$ the cross-sectional diameter of the Na atom for collisions of the first kind with electrons are known, where ϵ_{12} is the excitation energy of the 3^2P from the ground state 3^2S. By detailed balancing (**6**) we find that

$$\sigma_2^2(\epsilon)=\frac{q_1}{q_2}\,\frac{\epsilon+\epsilon_{12}}{\epsilon}\,\sigma_1^2(\epsilon+\epsilon_{12}), \qquad (1.2)$$

where q_1 is the statistical weight of the lower state, q_2 of the upper.

For an optical depth $\tau_0(5485\text{ A})=0{\cdot}48$ in the continuous spectrum, we find for a particular model of the solar photosphere (**14**), the following values of the temperature, total gas pressure and partial pressure due to free electrons:

$$\left.\begin{aligned} T &= 5838\text{ deg. K}\\ p &= 6{\cdot}9\times10^4\text{ dyne cm}^{-2}\\ p_e &= 17{\cdot}7\quad\text{dyne cm}^{-2}\end{aligned}\right\}\text{ at }\tau_0(5485\text{ A})=0{\cdot}48.$$

Not widely dissimilar values are given by other photospheric models for a corresponding optical depth. Then from Christoph's corrected experimental values of σ_1^2, converted to σ_2^2 by equation (1.2) and a numerical integration of the expression (1.1), the number of collisions of the second kind* experienced by a Na atom in the 3^2P state at this depth in the photosphere is $1{\cdot}52\times10^6\text{ sec}^{-1}$. The average life of the excited Na atom in so far as it is determined by collisions of the second kind is therefore $t_0=0{\cdot}66\times10^{-6}$ sec. But the average life as determined by spontaneous emission is $1/A_{21}$, so that the probability that the atom does not experience a collision of the second kind in this lifetime is $\exp(-1/A_{21}t_0)$. From Trumpy's determination of the oscillator strength of the Na *D*-lines (**17**), $A_{21}=0{\cdot}62\times10^8\text{ sec}^{-1}$, or the probability that the atom does not experience a collision of the second kind in the lifetime $1/A_{21}$ sec is 0·976. In other words only 2·4 per cent of the excited Na atoms are quenched by a collision of the second kind converting the quantum previously absorbed into kinetic energy of the free electrons. The percentage is in fact even smaller since the life before spontaneous emission is determined by the roughly ten times shorter life produced by collisional damping. On this basis it is readily found that only 0·16 per cent of the excited Na atoms give up their energy as a result of collisions of the second kind.

Since at this depth in the photosphere the hydrogen atoms are some $3{\cdot}9\times10^3$ times more abundant than the free electrons, it might be expected that the excited Na atom would experience many collisions of the second kind with H

* The number of the first kind is $7{\cdot}90\times10^4\text{ sec}^{-1}$.

atoms. As Massey and Burhop (**8**) point out, however, this neglects the fact that atoms approach one another relatively slowly; as a consequence such collisions are nearly adiabatic, the excited atom having a chance to adjust itself without a quantum transition taking place. In fact collisions with H atoms lead to a perturbation of the energy level and to collisional broadening, not to collisions of the second kind. The probability of a quenching collision will be small (**8**) if $\sigma\epsilon_{12}/hV \gg 1$, where σ is the gas-kinetic cross-sectional diameter of the atom and V the relative velocity. From Rosin and Rabi's value of $\sigma^2 = 41{\cdot}38 \times 10^{-16}\,\mathrm{cm}^2$ for the Na atom (**15**) and a mean velocity of H atoms of $1{\cdot}2 \times 10^6\,\mathrm{cm\,sec}^{-1}$, this quantity equals 27·2 and the encounters are nearly adiabatic.

A more direct comparison may be made with experiment. Norrish and Smith (**11**) have measured σ_2^2 for Na atoms in the 3^2P state in atmospheres of various monatomic rare gases and of many molecules. For an atmosphere of He they find, for example, that $\sigma_2^2 \simeq 0$, and for methane $= 0{\cdot}11 \times 10^{-16}\,\mathrm{cm}^2$. In an atmosphere of H atoms, as in the photosphere, the value would presumably be nearer that for a He than a methane atmosphere. Assume, however, that it is comparable with the latter and is independent of relative kinetic energy. Then from equation (1.1) the number of quenching collisions experienced per sec by an excited Na atom will be

$$2n_2\sigma_2^2\sqrt{2\pi kT\left(\frac{1}{m_1}+\frac{1}{m_2}\right)} = 3{\cdot}35 \times 10^8\,\mathrm{sec}^{-1}$$

due to encounters with H atoms at this depth in the photosphere. If, as in the previous discussion, the interval before spontaneous emission is determined by collisional broadening, we find that only 0·46 per cent of the excited Na atoms experience a collision of the second kind with H atoms.

1.2. *Runaway quanta.*—In so far as the behaviour of the Na atom is typical, we may conclude that collisions of the second kind play no significant part in converting the exciting light quantum then and there into kinetic energy of free electrons and H atoms. Nor, so far as I am aware, has it ever been suggested that they do. What has been suggested is that the originally absorbed quantum, or one of different frequency, is re-emitted and re-absorbed many times until a collision of the second kind or a photo-ionization takes place.

We are interested, however, in finding the source function at a particular place. In most observational investigations this "place" is somewhat narrowly defined, and consists of a region with a superficial area of perhaps five seconds of arc square, or 3600 km square, with an optical depth of 0·1 or smaller. The quantum, pictured as being handed about from atom to atom until a true absorption takes place, has a mean free path, as Unsöld (**18**) has pointed out in another connection, with an optical depth in the line absorption coefficient of the order of unity. Assuming that the probability of a collision of the second kind with free electrons and H atoms is of the order of 0·01, as indicated in the previous subsection, then after 100 transfers from atom to atom there will from the Poisson distribution be a probability of $1/e$ that a collision of the second kind has taken place. In its "drunkard's walk" of 100 paces, each of optical depth unity, the quantum will on the average be 10 units of optical depth away from the place where we are determining the source function before there is a probability $1/e$ that a collision of the second kind occurs. In short, as far as this particular region of the photosphere is concerned, the quantum has been scattered, not absorbed.

2. *The exact equation of transfer and the coefficients of absorption and emission*

2.1. *The equation of transfer.*—The Fraunhofer lines are usually regarded as formed in accordance with the Eddington–Milne equation of transfer

$$\cos\theta \frac{dI_\nu}{\rho dz} = -(\kappa_0 + \kappa_\nu)I_\nu + (1-\epsilon)\kappa_\nu \int I_\nu \frac{d\omega}{4\pi} + (\kappa_0 + \epsilon\kappa_\nu)B_\nu,$$

where ϵ is the fraction of excited atoms experiencing a collision of the second kind. Obviously the equation is much simplified when $\epsilon = 1$, but this implies true absorption at this particular level which, as we have seen in Section 1, is improbable. On the other hand, if $\epsilon \neq 1$, the equation implies that the scattering is strictly coherent—a physically impossible assumption which, from Houtgast's investigation (**7**), is not realized in the photosphere. Moreover the only quanta emitted in the frequency of the line are assumed to arise from a population of the upper state solely determined by excitation by quanta of the line under consideration.

There are therefore good grounds for preferring the exact equation of transfer (**13**) for line formation

$$\cos\theta \frac{dI_\nu}{\rho dz} = -(\kappa_0 + \kappa_\nu)I_\nu + j_0 + j_\nu, \tag{2.1}$$

where κ_0 and κ_ν are the mass coefficients of absorption in the continuum and the line, both for the frequency ν, and where j_0 and j_ν are the mass coefficients of emission for the frequency ν in continuum and line. The latter coefficients are defined as the isotropic flux emitted at the frequency ν per unit frequency per unit solid angle per unit mass. Attention is therefore always focused on the single frequency ν and no mechanism of absorption or emission is pre-supposed. Introducing the optical depth at the frequency ν in the continuum

$$\tau_0 = -\int^z \kappa_0 \rho \, dz, \tag{2.2}$$

the exact equation takes the form

$$\cos\theta \frac{dI_\nu(\tau_0,\theta)}{d\tau_0} = (1+\eta)I_\nu(\tau_0,\theta) - \frac{j_0}{\kappa_0} - \eta\frac{j_\nu}{\kappa_\nu}, \tag{2.3}$$

where the source functions, j_0/κ_0 in the continuum and j_ν/κ_ν in the line, and $\eta = \kappa_\nu/\kappa_0$ are each functions of the optical depth τ_0. Without some knowledge of these functions, however, equation (2.3) is numerically almost unintegrable. Accordingly in the remainder of this section an attempt will be made to evaluate these functions in terms of atomic transition coefficients.

2.2. *Atomic transition coefficients for continuous absorption and emission.*—Following Milne (**9**) the atomic transition coefficients for photo-ionization and recombination may be readily derived. The number of photo-ionizations from the rth state of the neutral atom per unit volume per unit time is

$$\frac{dn_r}{dt} = -n_r \phi_r(\nu) u_\nu \, d\nu, \tag{2.4}$$

where n_r is the number of such atoms per unit volume, $u_\nu \, d\nu$ is the radiative energy density and $\phi_r(\nu)$ is the atomic probability of photo-ionization. The velocity V of the ejected electron is given by

$$h\nu = \tfrac{1}{2}mV^2 - E_r, \tag{2.5}$$

where E_r is the (negative) energy of the rth state. Correspondingly the number of recombinations of V-electrons per unit volume per unit time is

$$\frac{dn_1'}{dt} = -4\pi^2 n_1' n_e \mu(V) V\, dV[\sigma_r^2(V) + s_r^2(V) u_\nu], \tag{2.6}$$

where n_1' is the number of ionized atoms in the ground state per unit volume and n_e is the number of free electrons with the correct spin per unit volume. $\mu(V)dV$ is the fraction of electrons moving with a velocity between V and $V+dV$ per unit solid angle, or

$$\mu(V)\,dV = (m/2\pi kT)^{3/2} \exp(-mV^2/2kT) V^2\, dV. \tag{2.7}$$

$\sigma_r(V)$ and $s_r(V)$ are the spontaneous and stimulated cross-sectional diameters of the ionized atom for capturing an electron moving with velocity V on the rth state. Then in thermodynamical equilibrium at temperature T we have detailed balancing so that $dn_r/dt = dn_1'/dt$. Solving for u_ν and recalling that it is given by Planck's law we find from equations (2.4) and (2.6)

$$u_\nu = \frac{\sigma_r^2(V)}{s_r^2(V)} \left[\frac{n_r}{4\pi^2 n_1' n_e \mu(V) V} \frac{\psi_r(\nu)}{s_r^2(V)} \frac{d\nu}{dV} - 1 \right]^{-1} = \frac{8\pi h\nu^3}{c^3} [\exp(h\nu/kT) - 1]^{-1}. \tag{2.8}$$

From the Boltzmann equation

$$\frac{n_r}{n} = \frac{q_r}{b(T)} \exp\{-(E_r - E_1)/kT\}, \quad \frac{n_1'}{n'} = \frac{q_1'}{b'(T)},$$

and from Saha's ionization equation

$$\frac{n'}{n} p_e = \frac{2(2\pi m)^{3/2}}{h^3} \frac{b'(T)}{b(T)} (kT)^{5/2} \exp(E_1/kT),$$

where $p_e = 2n_e kT$, since the electron has two possible orientations in an applied magnetic field. $b(T)$ and $b'(T)$ are the partition functions for the neutral and ionized atom respectively. Recalling from equation (2.5) that $d\nu/dV = mV/h$ we find after a little reduction the following relations between the transition coefficients for photo-ionization and recombination

$$\frac{\sigma_r^2(V)}{s_r^2(V)} = \frac{8\pi h\nu^3}{c^3}; \quad \frac{\sigma_r^2(V)}{\psi_r(\nu)} = \frac{q_r}{q_1'} \frac{2}{\pi m^2 V^2} \left(\frac{h\nu}{c}\right)^3. \tag{2.9}$$

Since these are strictly atomic constants, they and the relations between them remain unchanged no matter what the external conditions in which the atom finds itself. The atomic absorption coefficient for continuous radiation is thus always given by

$$\alpha_r(\nu) = (h\nu/c)\psi_r(\nu). \tag{2.10}$$

Consequently if N_r be the number of neutral atoms in state r per unit mass of the photosphere, the continuous mass coefficient of absorption due to these atoms is

$$\kappa_\theta = \frac{h\nu}{c} N_r \psi_r(\nu). \tag{2.11}$$

Similarly if N_1' is the number of ionized atoms in the ground state per unit mass of the photosphere, the mass coefficient of emission for the continuum is from equation (2.6)

$$j_\theta = \frac{h\nu}{4\pi} 4\pi^2 N_1' \frac{p_e}{2kT} \mu(V) V \sigma_r^2(V) \frac{dV}{d\nu}, \tag{2.12}$$

neglecting the stimulated transitions.

2.3. *The atomic line transition coefficients.*—As Spitzer (**16**), following Weisskopf and Wigner, has pointed out, it is statistically correct when dealing with an assemblage of atoms to regard absorption and emission lines as originating from transitions between broadened energy levels. The probability of an atom finding itself in a sub-level between s_r and s_r+ds_r, measured in frequency units, from the centre of the level is given (for atoms in an enclosure in thermodynamical equilibrium) by the normalized expression

$$\frac{\delta_r}{\pi}\frac{ds_r}{s_r^2+\delta_r^2}. \tag{2.13}$$

δ_r is the half half-width of the broadened energy level and is given by $\delta_r=1/4\pi t_r$, where t_r is the mean lifetime in state r as determined by radiative and collisional damping. The atomic probability of transition between a particular sub-level, s_1, of the ground state and a level between s_2 and s_2+ds_2 of state 2 is

$$B_{12}\frac{\delta_2}{\pi}\frac{ds_2}{s_2^2+\delta_2^2}. \tag{2.14}$$

The spontaneous and stimulated downward transition probabilities from a particular sub-level, s_2, to a sub-level between s_1 and s_1+ds_1 are likewise given by

$$A_{21}\frac{\delta_1}{\pi}\frac{ds_1}{s_1^2+\delta_1^2}, \quad B_{21}\frac{\delta_1}{\pi}\frac{ds_1}{s_1^2+\delta_1^2}. \tag{2.15}$$

When the atoms are not in an enclosure in thermodynamic equilibrium, these transition probabilities, since they are purely atomic quantities, remain unchanged. On the other hand, though the potentially realizable sub-levels are still in existence, the actual distribution of atoms among these sub-levels is no longer given by expression (2.13). Under these circumstances let the actual probability in the ground state be given by the normalized expression $W(s_1)\,ds_1$. Then if N_1^* is the total number of metal atoms in state 1 per unit mass of photospheric matter, the mass coefficient of absorption for atoms at rest is given by

$$\frac{h\nu_{12}^0}{c}N_1^*B_{12}\frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{s_2^2+\delta_2^2}=\frac{h\nu_{12}^0}{c}N_1^*B_{12}\frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{(s_1+\Delta\nu)^2+\delta_2^2}. \tag{2.16}$$

In these expressions ν_{12}^0 is the constant frequency for a transition between the centres of the two broadened states, and the integrations are to be carried out subject to the condition that $\nu_{12}=\nu_{12}^0+\Delta\nu$ is a constant.

The atoms are actually in motion, and for an atom with a velocity component, u, in the line of sight the central frequency of the absorbed line is $\nu_{12}^0-u(\nu_{12}^0/c)$. The fraction of atoms with velocity components between u and $u+du$ is

$$(m^*/2\pi kT_k)^{1/2}\exp(-m^*u^2/2kT_k)\,du,$$

where T_k is the kinetic temperature and m^* the mass of the metal atom. From the expression in (2.16) the mass coefficient of absorption for the frequency $\nu_{12}^0+\Delta\nu$ will clearly be given by

$$\kappa(\nu_{12}^0+\Delta\nu)=\frac{h\nu_{12}^0}{c}N_1^*B_{12}\left(\frac{m^*}{2\pi kT_k}\right)^{1/2}\frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\exp(-m^*u^2/2kT_k)\,du$$
$$\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{[s_1+\Delta\nu+u(\nu_{12}^0/c)]^2+\delta_2^2}. \tag{2.17}$$

Substituting

$$u=y\sqrt{\frac{2kT_k}{m^*}}, \quad b=\frac{\nu_{12}^0}{c}\sqrt{\frac{2kT_k}{m^*}}, \tag{2.18}$$

where b is a species of half width for Doppler broadening, and

$$\sigma = (s_1 + \Delta\nu)/b, \tag{2.19}$$

equation (2.17) becomes, inverting order of integration,

$$\kappa(\nu_{12}{}^0 + \Delta\nu) = \frac{h\nu_{12}{}^0}{c} N_1{}^* B_{12} \frac{\delta_2}{b^2\pi^{3/2}} \int_{-\infty}^{\infty} W(s_1)\, ds_1 \int_{-\infty}^{\infty} \frac{\exp(-y^2)\, dy}{(\sigma + y)^2 + (\delta_2/b)^2}, \tag{2.20}$$

where σ is a constant for the first integration.

The first integral in equation (2.20) is well known (**10**), and provided δ_2/b is small (under photospheric conditions it is of the order of $1{\cdot}4 \times 10^{-2}$) we may write

$$\int_{-\infty}^{\infty} \frac{\exp(-y^2)\, dy}{(\sigma + y)^2 + (\delta_2/b)^2} = \frac{\pi b}{\delta_2} \exp(-\sigma^2) - 2\sqrt{\pi}\{1 - 2\sigma F(\sigma)\}, \tag{2.21}$$

where $F(\sigma) = \exp(-\sigma^2) \int_0^{\sigma} \exp(x^2)\, dx$ is a tabulated function. Hence from equations (2.20), (2.21) and (2.19), the mass coefficient of absorption takes the form

$$\kappa(\nu_{12}{}^0 + \Delta\nu) = \frac{h\nu_{12}{}^0}{c} N_1{}^* B_{12} \frac{\delta_2}{b^2\pi^{3/2}} \int_{-\infty}^{\infty} \left[\frac{\pi b}{\delta_2} \exp\left\{ -\left(\frac{s_1 + \Delta\nu}{b} \right)^2 \right\} - 2\sqrt{\pi} \left\{ 1 - 2\frac{s_1 + \Delta\nu}{b} F\left(\frac{s_1 + \Delta\nu}{b} \right) \right\} \right] W(s_1)\, ds_1. \tag{2.22}$$

The integral in equation (2.22) can be written

$$\int_{-\infty}^{\infty} \frac{\pi b}{\delta_2} \exp\left\{ -\left(\frac{s_1 + \Delta\nu}{b} \right)^2 \right\} W(s_1)\, ds_1 - \int_{-\infty}^{\infty} 2\sqrt{\pi} \left\{ 1 - 2\frac{s_1 + \Delta\nu}{b} F\left(\frac{s_1 + \Delta\nu}{b} \right) \right\} W(s_1) ds_1$$
$$\simeq \frac{\pi b}{\delta_2} \exp\left\{ -\left(\frac{\Delta\nu}{b} \right)^2 \right\} - 2\sqrt{\pi} \left\{ 1 - \frac{2\Delta\nu}{b} F\left(\frac{\Delta\nu}{b} \right) \right\},$$

where, in order to estimate the relative magnitude of the two integrals, we have assumed $W(s_1)$ to be a delta function. Then from the tabulated values of the $F(\Delta\nu/b)$ function, a value of $b = 3{\cdot}77 \times 10^9\,\mathrm{sec}^{-1}$ ($= 0{\cdot}044$ A corresponding to a Na atom at 6000 deg. K) and $\delta_2 = 5{\cdot}08 \times 10^7\,\mathrm{sec}^{-1}$ ($= 0{\cdot}00059$ A) we find for $\Delta\nu \leqslant 1{\cdot}4b$ that the first term is at least 33·3 times greater than the second, and for $\Delta\nu \geqslant 3{\cdot}4b$ that the second term is at least 81·2 times greater than the first. Further for large values of $\Delta\nu/b$

$$1 - 2\frac{\Delta\nu}{b} F\left(\frac{\Delta\nu}{b} \right) = -\left[\frac{1}{2(\Delta\nu/b)^2} + \frac{3}{\{2(\Delta\nu/b)^2\}^2} + \cdots \right],$$

so that for $\Delta\nu \geqslant 6b$ the error in putting this expression equal to $-\frac{1}{2}(\Delta\nu/b)^2$ is less than 4 per cent. Hence with an error not in excess of this amount we may write equation (2.22) for the mass coefficient of absorption

$$\Delta\nu \geqslant 6b \qquad \kappa(\nu_{12}{}^0 + \Delta\nu) = \frac{h\nu_{12}{}^0}{c} N_1{}^* B_{12} \frac{\delta_2}{\pi} \int_{-\infty}^{\infty} \frac{W(s_1)\, ds_1}{(s_1 + \Delta\nu)^2}, \tag{2.23}$$

$$\Delta\nu \leqslant 1{\cdot}4b \qquad = \frac{h\nu_{12}{}^0}{c} N_1{}^* B_{12} \frac{1}{b\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left\{ -\left(\frac{s_1 + \Delta\nu}{b} \right)^2 \right\} W(s_1)\, ds_1. \tag{2.24}$$

The expression for the mass coefficient of emission may be derived similarly. Where $N_2{}^*$ is the total number of metal atoms in state 2 per unit mass of photospheric matter and where we neglect stimulated downward transitions we find for

$$\Delta\nu \geqslant 6b \quad j(\nu_{12}{}^0 + \Delta\nu) = \frac{h\nu_{12}{}^0}{4\pi} N_2{}^* A_{21} \frac{\delta_1}{\pi} \int_{-\infty}^{\infty} \frac{W(s_2)\, ds_2}{(s_2 - \Delta\nu)^2} \tag{2.25}$$

$$\Delta\nu \leqslant 1{\cdot}4b \qquad = \frac{h\nu_{12}{}^0}{4\pi} N_2{}^* A_{21} \frac{1}{b\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left\{ -\left(\frac{s_2 - \Delta\nu}{b} \right)^2 \right\} W(s_2)\, ds_2. \tag{2.26}$$

2.4. *Assumed photospheric conditions.*—In deriving the values of the source functions we shall neglect the radiatively stimulated emissions. Thus in thermodynamical equilibrium the source function, given under these conditions by Kirchhoff's law, will be represented by Wien's law—a sufficiently accurate approximation to Planck's law for the optical part of the spectrum and photospheric temperatures. Further, we shall suppose that in the photosphere the electrons and atoms have a Maxwellian distribution of velocities corresponding to a common kinetic temperature T_k, that there is a Boltzmann distribution among the stationary states corresponding to an excitation temperature T_e for hydrogen and T_e^* for the metal, and that the degree of ionization may be represented by Saha's expression with an ionization temperature T_i for hydrogen and T_i^* for the metal.

As is customary we shall assume that the continuous absorption and emission arise from photo-ionization of the negative hydrogen ion, and its re-formation by the recombination of a free electron with the neutral hydrogen atom in its ground state. Free–free transitions, which contribute some 10 per cent of the continuous absorption and emission for $\lambda < 7000$ A, are for simplicity neglected, though, following Cillié (**5**), their effect could be readily included.

If n' be the total number of neutral hydrogen atoms per unit volume, the number, n_1', in the ground state is

$$n_1' = n' q_1' / b'(T_e). \tag{2.27}$$

The number of negative hydrogen ions per unit volume is

$$n = n' \phi(T_i) p_e, \tag{2.28}$$

where

$$\frac{1}{\phi(T_i)} = \frac{2(2\pi m)^{3/2}}{h^3} \frac{b'(T_e)}{b(T_e)} (kT_i)^{3/2} \exp(E_1/kT_i). \tag{2.29}$$

E_1 is the (negative) dissociation energy for the negative hydrogen ion. Formally the number of negative hydrogen ions in the ground state per unit volume is

$$n_1 = n q_1 / b(T_e), \tag{2.30}$$

though since there is only one state, and it single, $q_1 = b(T_e) = 1$. Hence from equations (2.27), (2.28) and (2.30)

$$n_1 = n_1' \frac{q_1}{q_1'} \frac{b'(T_e)}{b(T_e)} \phi(T_i) p_e. \tag{2.31}$$

Assuming the photosphere to be composed overwhelmingly of neutral hydrogen in its ground state, the number of hydrogen atoms in the ground state per unit mass of photospheric matter is

$$N_1' = 1/m_H, \tag{2.32}$$

where m_H is the mass in gm of the hydrogen atom. Hence from equations (2.31) and (2.32) the number of negative hydrogen ions per gm of photospheric matter is

$$N_1 = \frac{1}{m_H} \frac{q_1}{q_1'} \frac{b'(T_e)}{b(T_e)} \phi(T_i) p_e. \tag{2.33}$$

N_1 and N_1' are the quantities to be inserted in equations (2.11) and (2.12) for the mass coefficients of continuous absorption and emission.

Again, if n^* be the total number per unit volume of neutral atoms of the metal under consideration, the numbers in states 1 and 2 will be

$$n_1^* = n^* q_1^* / b^*(T_e^*), \quad n_2^* = n^* q_2^* \exp\{-(E_2^* - E_1^*)/kT_e^*\} / b^*(T_e^*), \tag{2.34}$$

where E_1^* and E_2^* are the (negative) energies of these states. If $n^{*\prime}$ is the number of once ionized metal atoms in all states of excitation per unit volume we have

$$n^* = n^{*\prime}\phi^*(T_i^*)p_e, \tag{2.35}$$

where $\phi^*(T_i^*)$ is given by the expression corresponding to equation (2.29).

If the ratio of abundance of hydrogen to the specific metal be $A^* \simeq 10^4$, then

$$n^* + n^{*\prime} = (n + n')/A^*.$$

But under photospheric conditions $n \ll n'$ and $n^* \ll n^{*\prime}$, so that

$$n^{*\prime} \simeq n'/A^* = n_1' b'(T_e)/q_1' A^* \tag{2.36}$$

from equation (2.27). Hence

$$N_1^* = \frac{1}{m_H A^*} \frac{q_1^*}{b^*(T_e^*)} \frac{b'(T_e)}{q_1'} \phi^*(T_i^*) p_e, \tag{2.37}$$

$$N_2^* = \frac{1}{m_H A^*} \frac{q_2^* \exp\{-(E_2^* - E_1^*)/kT_e^*\}}{b^*(T_e^*)} \frac{b'(T_e)}{q_1'} \phi^*(T_i^*) p_e, \tag{2.38}$$

and these are the quantities to be inserted in equations (2.23) to (2.26) for the mass coefficients of line absorption and emission.

2.5. *Source function for the continuum.*—From equations (2.11), (2.12), (2.32) and (2.33) we find

$$\frac{j_0}{\kappa_0} = \frac{c}{4\pi}\left[\frac{4\pi^2}{2kT_k}\mu(V)V\right]\left[\frac{q_1 b'(T_e)}{q_1' b(T_e)}\phi(T_i)\right]^{-1}\frac{\sigma_1^2(V)}{\psi_1(\nu)}\frac{dV}{d\nu}.$$

In this expression $\mu(V)$ is given by equation (2.7) with the kinetic temperature T_k, $\phi(T_i)$ by equation (2.29), the ratio of atomic constants by equation (2.9), and from equation (2.5) $dV/d\nu = h/mV$. With these values we find

$$\frac{j_0}{\kappa_0} = \frac{2h\nu^3}{c^2}\left(\frac{kT_i}{kT_k}\right)^{5/2}\exp\left(-\frac{mV^2}{2kT_k} + \frac{E_1}{kT_i}\right). \tag{2.39}$$

In an enclosure containing hydrogen in thermodynamic equilibrium where $T_k = T_i = T$, equation (2.39) reduces, when account is taken of equation (2.5), to

$$\frac{j_0}{\kappa_0} = \frac{2h\nu^3}{c^2}\exp(-h\nu/kT), \tag{2.40}$$

that is to Wien's law as it should.

2.6. *Source function for the line.*—From equations (2.23), (2.25), (2.37) and (2.38) we find for $\Delta\nu \geq 6b$

$$\frac{j(\nu_{12}^0 + \Delta\nu)}{\kappa(\nu_{12}^0 + \Delta\nu)} = \frac{c}{4\pi}\frac{q_2^*\exp\{-(E_2^* - E_1^*)/kT_e^*\}}{q_1^*}\frac{A_{21}}{B_{12}}$$
$$\left[\frac{\delta_1}{\pi}\int_{-\infty}^{\infty}\frac{W(s_2)\,ds_2}{(s_2 - \Delta\nu)^2}\Big/\frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{(s_1 + \Delta\nu)^2}\right].$$

From Bohr's frequency condition $E_2^* - E_1^* = h\nu_{12}^0$, while the relation of the Einstein coefficients, defined for energy density, is

$$\frac{A_{21}}{B_{12}} = \frac{q_1^*}{q_2^*}\frac{8\pi h(\nu_{12}^0)^3}{c^3}.$$

Hence

$$\frac{j(\nu_{12}^0 + \Delta\nu)}{\kappa(\nu_{12}^0 + \Delta\nu)} = \frac{2h(\nu_{12}^0)^3}{c^2}\exp(-h\nu_{12}^0/kT_e^*)\left[\frac{\delta_1}{\pi}\int_{-\infty}^{\infty}\frac{W(s_2)\,ds_2}{(s_2 - \Delta\nu)^2}\Big/\frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{(s_1 + \Delta\nu)^2}\right]. \tag{2.41}$$

Similarly for $\Delta\nu \leqslant 1{\cdot}4b$, we have from equations (2.24) and (2.26)

$$\frac{j(\nu_{12}{}^0+\Delta\nu)}{\kappa(\nu_{12}{}^0+\Delta\nu)} = \frac{2h(\nu_{12}{}^0)^3}{c^2}\exp(-h\nu_{12}{}^0/kT_e{}^*)$$
$$\times\left[\int_{-\infty}^{\infty}\exp\left\{-\left(\frac{s_2-\Delta\nu}{b}\right)^2\right\}W(s_2)\,ds_2\Big/\int_{-\infty}^{\infty}\exp\left\{-\left(\frac{s_1+\Delta\nu}{b}\right)^2\right\}W(s_1)\,ds_1\right]. \quad (2.42)$$

In an enclosure in thermodynamic equilibrium containing the metal atoms only, $T_e{}^* = T$ and the expressions $W(s)\,ds$ are given by (2.13). Hence in equation (2.41) we may write the numerator of the expression in square brackets:

$$\frac{\delta_1}{\pi}\int_{-\infty}^{\infty}\frac{W(s_2)\,ds_2}{(s_2-\Delta\nu)^2} \simeq \frac{\delta_1}{\pi}\int_{-\infty}^{\infty}\frac{W(s_2)\,ds_2}{(s_2-\Delta\nu)^2+\delta_1{}^2} = \frac{\delta_1\delta_2}{\pi^2}\int_{-\infty}^{\infty}\frac{ds_2}{[(s_2-\Delta\nu)^2+\delta_1{}^2][s_2{}^2+\delta_2{}^2]}$$
$$= \frac{\delta_1+\delta_2}{\pi}\frac{1}{(\Delta\nu)^2+(\delta_1+\delta_2)^2},$$

while the denominator similarly has an identical value. Hence in thermodynamic equilibrium equation (2.41) reduces, as it should, to Wien's law. Again, in equation (2.42) the numerator of the expression in square brackets becomes

$$\int_{-\infty}^{\infty}\exp\left\{-\left(\frac{s_2-\Delta\nu}{b}\right)^2\right\}W(s_2)\,ds_2 = \frac{\delta_2}{\pi}\int_{-\infty}^{\infty}\left[\exp\left\{-\left(\frac{s_2-\Delta\nu}{b}\right)^2\right\}\Big/(s_2{}^2+\delta_2{}^2)\right]ds_2$$
$$= \frac{\delta_2}{\pi b}\left[\frac{\pi b}{\delta_2}\exp\left\{-\left(\frac{\Delta\nu}{b}\right)^2\right\} - 2\sqrt{\pi}\left\{1-\frac{2\Delta\nu}{b}F\left(\frac{\Delta\nu}{b}\right)\right\}\right] \simeq \exp\left\{-\left(\frac{\Delta\nu}{b}\right)^2\right\},$$

where the integration is effected as in equation (2.21) and the last step follows since $\Delta\nu \leqslant 1{\cdot}4b$. The denominator of the expression in square brackets has similarly an identical value, so that in thermodynamical equilibrium equation (2.42) also reduces to Wien's law.

A further test for the accuracy of our expressions for the mass coefficients of absorption and emission in the continuum and line spectrum may readily be applied. Suppose we had a sample of our photospheric matter with hydrogen and the fraction $1/A^*$ of the metal in an enclosure in thermodynamical equilibrium at some temperature comparable with that of the photosphere. Then in the frequency of the line the mass coefficient of emission will be $j_0 + j(\nu_{12}{}^0+\Delta\nu)$, and the mass coefficient of absorption $\kappa_0 + \kappa_2(\nu_{12}{}^0+\Delta\nu)$. Inserting the values of these coefficients as given above and remembering that in thermodynamical equilibrium $T_k = T_e = T_e{}^* = T_i = T_i{}^* = T$, we find after a little reduction that

$$\frac{j_0 + j(\nu_{12}{}^0+\Delta\nu)}{\kappa_0 + \kappa(\nu_{12}{}^0+\Delta\nu)} = \frac{2h\nu^3}{c^3}\exp(-h\nu/kT),$$

reducing to Wien's law as it should.

2.7. *The ratio of absorption coefficients.*—For $\Delta\nu \geqslant 6b$ we find from equations (2.11), (2.23), (2.33) and (2.37), since $\nu = \nu_{12}{}^0$,

$$\eta = \frac{\kappa(\nu_{12}{}^0+\Delta\nu)}{\kappa_0} = \left[\frac{1}{A^*}\frac{q_1{}^*}{b^*(T_e{}^*)}\frac{b'(T_e)}{q_1{}'}\phi^*(T_i{}^*)B_{12}\frac{\delta_2}{\pi}\right.$$
$$\left.\times\int_{-\infty}^{\infty}\frac{W(s_1)\,ds_1}{(s_1+\Delta\nu)^2}\right]\left[\frac{q_1}{b(T_e)}\frac{b'(T_e)}{q_1{}'}\phi(T_i)\psi_1(\nu)\right]^{-1}.$$

$\phi(T_i)$ is given by equation (2.29) and $\phi^*(T_i{}^*)$ by

$$\frac{1}{\phi^*(T_i{}^*)} = \frac{2(2\pi m)^{3/2}}{h^3}\frac{b^{*\prime}(T_e{}^*)}{b^*(T_e)}(kT_i{}^*)^{3/2}\exp(E_1{}^*/kT_i{}^*),$$

where $-E_1^*$ is the ionization potential of the neutral metal. Inserting these values we have

$$\eta = \left[\frac{1}{A^*}\frac{q_1^*}{q_1}\frac{b'(T_e)}{b^{*\prime}(T_e^*)}\frac{B_{12}}{\psi_1(\nu)}\right]\left(\frac{kT_i}{kT_i^*}\right)^{3/2}\exp\left(\frac{E_1}{kT_i}-\frac{E_1^*}{kT_i^*}\right)\int_{-\infty}^{\infty}\frac{\delta_2}{\pi}\frac{W(s_1)\,ds_1}{(s_1+\Delta\nu)^2}. \quad (2.43)$$

For $\Delta\nu \leqslant 1{\cdot}4b$ we find from equation (2.24) an expression identical with the right-hand side of equation (2.43), except that the integral is replaced by

$$\int_{-\infty}^{\infty}\frac{1}{b\sqrt{\pi}}\exp\left\{-\left(\frac{s_1+\Delta\nu}{b}\right)^2\right\}W(s_1)\,ds_1. \quad (2.44)$$

3. *Numerical integration of exact equation of transfer*

Multiplying through by the integrating factor

$$R = R_0\exp\left\{-\sec\theta\int_0^{\tau_0}(1+\eta)\,dt\right\}$$

the exact equation of transfer (2.3) may be written

$$I_\nu(0,\theta) = \int_0^{\infty}\left(\frac{j_0}{\kappa_0}+\eta\frac{j_\nu}{\kappa_\nu}\right)\exp\left\{-\sec\theta\int_0^{\tau_0}(1+\eta)\,dt\right\}\sec\theta\,d\tau_0. \quad (3.1)$$

This is an integral equation from which to find $\frac{j_0}{\kappa_0}$, $\frac{j_\nu}{\kappa_\nu}$ and η as functions of optical depth τ_0 (in the frequency of the adjacent continuous spectrum) from the observed limb darkening in the line. Its numerical integration may be effected from the knowledge we have gained in the previous section of the dependence of these three functions on kinetic (T_k), ionization (T_i and T_i^*) and excitation (T_e and T_e^*) temperatures. In order to put these functional relations in a more tractable form, three simplifying assumptions may be made.

The first assumption is that $T_k = T_i = T_r$, where T_r is the radiation temperature. The justification for this assumption has been given elsewhere (**14**). If it is legitimate we have from equations (2.39) and (2.5)

$$\frac{j_0}{\kappa_0} = \frac{2h\nu^3}{c^2}\exp(-h\nu/kT_r). \quad (3.2)$$

The resulting value of T_r as a function of τ_0 can be found in one of the usual ways from the numerical integration of equation (3.1) for the case $\eta = 0$, that is from the darkening to the limb, $I_0(0,\theta)$ in the adjacent continuous spectrum.

The second and more dubious assumption is that T_i^* (for the metal) $= T_r$. This assumption is questionable first because the frequency of the ionizing radiation, at least for ionization from the ground state of the metal, is far removed from the frequency ν for which $T_r(\nu)$ is determined as a function of τ_0 from equation (3.2). Secondly, since the radiation field is crossed by the already formed absorption lines of the metal, the excitation temperature, T_e^*, must be less than T_r so that T_i^* must be less than T_r. The effect on the degree of ionization may however be small (**14**). The consequence of this simplifying assumption is that equation (2.43) takes the form

$$\eta = \left[\frac{1}{A^*}\frac{2q_1^*}{b^{*\prime}(T_e^*)}\frac{B_{12}}{\psi_1(\nu)}\right]\exp\{(E_1-E_1^*)/kT_r\}\int_{-\infty}^{\infty}\frac{\delta_2}{\pi}\frac{W(s_1)\,ds_1}{(s_1+\Delta\nu)^2}, \quad (3.3)$$

where for the negative hydrogen ion $q_1 = 1$ and for the neutral atom $b'(T_e) = q_1' = 2$, since all the neutral hydrogen atoms are effectively in the ground state. Equation

(3.3) applies to the case where $\Delta\nu \geqslant 6b$; for $\Delta\nu \leqslant 1{\cdot}4b$ the equation remains the same except that the integral is replaced by that in (2.44).

Such integrals containing the unknown functions $W(s_1)$ and $W(s_2)$ involve us in a certain difficulty. Ideally we should like to determine these functions of s and τ_0, but the Fraunhofer lines by themselves do not appear to yield sufficient information. Applying the mean value theorem

$$\frac{\delta_r}{\pi}\int_{-\infty}^{\infty}\frac{W(s_r)\,ds_r}{(s_r+\Delta\nu)^2}=\frac{\delta_r}{\pi}\frac{1}{(s_r'+\Delta\nu)^2}\int_{-\infty}^{\infty}W(s_r)\,ds_r=\frac{\delta_r}{\pi}\frac{1}{(s_r'+\Delta\nu)^2},$$

the last step following since $W(s_r)$ is a normalized function. In this expression s_r' is a particular value of s_r.

Our third assumption is that since $W(s_r)$ is a peaked distribution centred on $s_r=0$ we may put $s_r'=0$. In thermodynamical equilibrium $W(s_r)$ is given by the expression (2.13), which is very peaked indeed, and in the photosphere will differ from this only in that not all the sub-levels will be filled; it may however be expected still to be peaked. With this assumption equation (3.3) for η takes the form

$$\Delta\nu \geqslant 6b \qquad \eta=\frac{1}{(\Delta\nu)^2}\left[\frac{1}{\pi A^*}\frac{2q_1^*}{b^{*\prime}(T_e^*)}\frac{B_{12}}{\psi_1(\nu)}\right]\delta_2\exp\{(E_1-E_1^*)/kT_r\}, \tag{3.4}$$

$$\Delta\nu \leqslant 1{\cdot}4b \qquad =\exp\left\{-\left(\frac{\Delta\nu}{b}\right)^2\right\}\left[\frac{1}{A^*\sqrt{\pi}}\frac{2q_1^*}{b^{*\prime}(T_e^*)}\frac{B_{12}}{\psi_1(\nu)}\right]\frac{1}{b}\exp\{(E_1-E_1^*)/kT_r\}. \tag{3.5}$$

In these expressions the quantity in square brackets is independent both of τ_0 and of $\Delta\nu$. Apart from the unknown abundance ratio of the metal, which we wish to find, it involves only known atomic constants. In front of the square brackets stands the factor which depends primarily on $\Delta\nu$ and after stand the factors which depend upon T_r, that is upon the optical depth τ_0. Of these latter factors the dominant one is the exponential function which is completely known, $-E_1$ being the dissociation potential of the negative hydrogen ion, $-E_1^*$ the ionization potential of the metal while T_r is known as a function of τ_0 from equation (3.2). Of the remaining factors dependent upon optical depth, b may be calculated from equation (2.18) and δ_2 from the theory of collisional broadening, since both the gas pressure, p, and $T_k=T_r$ are known (**14**) as functions of optical depth, τ_0. We may therefore write equations (3.4) and (3.5)

$$\Delta\nu \geqslant 6b \qquad \eta=\frac{C_1}{(\Delta\nu)^2}f_1(\tau_0), \tag{3.6}$$

$$\Delta\nu \leqslant 1{\cdot}4b \qquad =C_2 f_2(\tau_0)\exp\left[-\{\Delta\nu/g(\tau_0)\}^2\right], \tag{3.7}$$

where $f_1(\tau_0)$, $f_2(\tau_0)$ and $g(\tau_0)$ are known functions of τ_0, and C_1 and C_2 are constants to be determined.

Finally and by virtue of our third assumption we may write equations (2.41) and (2.42) in the form

$$\Delta\nu \geqslant 6b \qquad \frac{j(\nu_{12}^0+\Delta\nu)}{\kappa(\nu_{12}^0+\Delta\nu)}=\frac{\delta_1}{\delta_2}\frac{2h(\nu_{12}^0)^3}{c^2}\exp(-h\nu_{12}^0/kT_e^*), \tag{3.8}$$

$$\Delta\nu \leqslant 1{\cdot}4b \qquad =\frac{2h(\nu_{12}^0)^3}{c^2}\exp(-h\nu_{12}^0/kT_e^*). \tag{3.9}$$

Following a suggestion originally made by ten Bruggencate (**3**), the numerical solution of the integral equation (3.1) may now be made. By a numerical solution of this equation for $\eta=0$ we find j_0/κ_0, therefore from equation (3.2) T_r, as a

function of τ_0. Then in a specific frequency in the line, namely $\nu_{12} = \nu_{12}^0 + \Delta\nu_1$ for the case $\Delta\nu \geq 6b$, arbitrarily choose some value C_1' of the constant in equation (3.6). From this equation we then know η as an explicit function of τ_0, and we may therefore solve equation (3.1) to find the only unknown function of optical depth, that is $j'(\nu_{12}^0 + \Delta\nu_1)/\kappa'(\nu_{12}^0 + \Delta\nu_1)$. At some second frequency $\Delta\nu_2$, the corresponding values of η as a function of optical depth are from equation (3.6) given by

$$\eta_2'(\tau_0) = \left(\frac{\Delta\nu_1}{\Delta\nu_2}\right)^2 \eta_1'(\tau_0).$$

Now solve equation (3.1) for $j'(\nu_{12}^0 + \Delta\nu_2)/\kappa'(\nu_{12}^0 + \Delta\nu_2)$. Then if we have chosen our constant C_1' correctly, since from equation (3.8) $j(\nu_{12}^0 + \Delta\nu)/\kappa(\nu_{12}^0 + \Delta\nu)$ is independent of $\Delta\nu$, the two thus found values of this function should be identical at each optical depth τ_0. If they are not another value C_1'' is selected and the whole calculation repeated, if necessary again and again, until the equality

$$\frac{j(\nu_{12}^0 + \Delta\nu_1)}{\kappa(\nu_{12}^0 + \Delta\nu_1)} = \frac{j(\nu_{12}^0 + \Delta\nu_2)}{\kappa(\nu_{12}^0 + \Delta\nu_2)}$$

is satisfied at each optical depth.

This procedure may be contrasted with that employed earlier (**12**) in the numerical integration of the Eddington–Milne equation. In that investigation η was an unknown *function* of τ_0, and it was only as a result of very protracted trials that its functional form was finally derived. In the present method, on the other hand, only a single constant, C_1, has to be found and a strictly iterative procedure may be followed. A further disadvantage of the earlier method, quite apart from the fact that it was applied to the inexact Eddington–Milne equation, was that the test for the correct functional form of η could only be applied after a tedious calculation of

$$\int I_\nu(\tau_0, \theta) \frac{d\omega}{4\pi}$$

at each optical depth—a calculation which took longer than the preceding solution of the integral equation. In the present method, on the other hand, the test is immediate and involves no further calculation after the solution of the integral equation for the frequencies $\Delta\nu_1$ and $\Delta\nu_2$. As far as can be judged without an actual trial the present method should therefore prove practicable and fairly rapid.

From the solution of the integral equation two results follow. First from the value of the constant C_1, which contains as unknown only the quantity A^* as is seen from equations (3.6) and (3.4), we obtain a direct observational determination of the abundance ratio of the metal to hydrogen. Secondly from the determination of $j(\nu_{12}^0 + \Delta\nu)/\kappa(\nu_{12}^0 + \Delta\nu)$ as a function of τ_0 we find from equation (3.8), and a calculated value of the ratio δ_1/δ_2, the value of the excitation temperature, that is the ratio of populations in states 2 and 1 of the metal atom, as a function of τ_0. If in addition the integral equation can be solved for the case $\Delta\nu \leq 1{\cdot}4b$ we obtain from the constant C_2 in equations (3.7) and (3.5) an independent determination of the abundance ratio, a direct determination of the excitation temperature from equation (3.9) and, combined with the already determined equation (3.8), an observed value of the ratio δ_1/δ_2. The solution for the case $\Delta\nu \leq 1{\cdot}4b$ follows with obvious slight modifications the method already outlined for $\Delta\nu \geq 6b$, but depends for its success upon an accurately

determined line core, an observationally difficult problem, and upon a reasonable change in surface brightness within $\Delta v \leqslant 1{\cdot}4b$. If the observational difficulties can be overcome it is probably the better method for faint lines, and the more directly informative.

The solution of the integral equation has been treated for the case of resonance lines, but it is not in principle limited to them. The only requirement for the application to any other line is that the population in the lower state of the line has been previously found relative to the ground state of the atom by the prior application of the method to an appropriate line.

In conclusion I wish to express my thanks to Mr R. J. Bray for many helpful discussions on the problem of line formation.

University Observatory,
Oxford :
1955 *March* 16.

References

(1) D. R. Bates, A. Fundaminsky, J. W. Leech and H. S. W. Massey, *Phil. Trans.* A, **243**, 121, 1950.
(2) K.-H. Böhm, *Zs. f. Ap.*, **35**, 179, 1954.
(3) P. ten Bruggencate, *Zs. f. Ap.*, **4**, 159, 1932.
(4) W. Cristoph, *Ann. d. Phys.*, **23**, 51, 1935.
(5) G. G. Cillié, *M.N.*, **92**, 820, 1932.
(6) R. H. Fowler, *Phil. Mag.*, **47**, 257, 1924.
(7) J. Houtgast, Dissertation, Utrecht, 1942.
(8) H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena*, p. 441, Oxford, 1952.
(9) E. A. Milne, *Phil. Mag.*, **47**, 209, 1924.
(10) A. C. G. Mitchell and M. W. Zemansky, *Resonance Radiation and Excited Atoms*, p. 321, Cambridge, 1934.
(11) R. G. W. Norrish and W. M. Smith, *Proc Roy. Soc.* A, **176**, 295, 1940.
(12) H. H. Plaskett, *M.N.*, **101**, 3, 1941; *Comm Univ. Obs., Oxford*, No. 19.
(13) H. H. Plaskett, *M.N.*, **107**, 146, 1947.
(14) H. H. Plaskett, *Physics of the Photosphere, Vistas in Astronomy*, I, p. 637, London, 1955.
(15) S. Rosin and I. I. Rabi, *Phys. Rev.*, **48**, 373, 1935.
(16) L. Spitzer, *M.N.*, **96**, 794, 1936.
(17) B. Trumpy, *Zs. f. Phys.*, **71**, 720, 1931.
(18) A. Unsöld, *Physik der Sternatmosphären*, p. 272, Berlin, 1938.
(19) A. Unsöld, *Zs. f. Ap.*, **21**, 22, 1941.
(20) A. Unsöld, *Ann. d. Phys.*, **3**, 124, 1948.

AN INTERFEROMETRIC INVESTIGATION OF THE EFFECT OF HYPERFINE STRUCTURE ON SOLAR LINE PROFILES

W. R. Hindmarsh

(Communicated by the Director, University Observatory, Oxford)

(Received 1955 March 23)

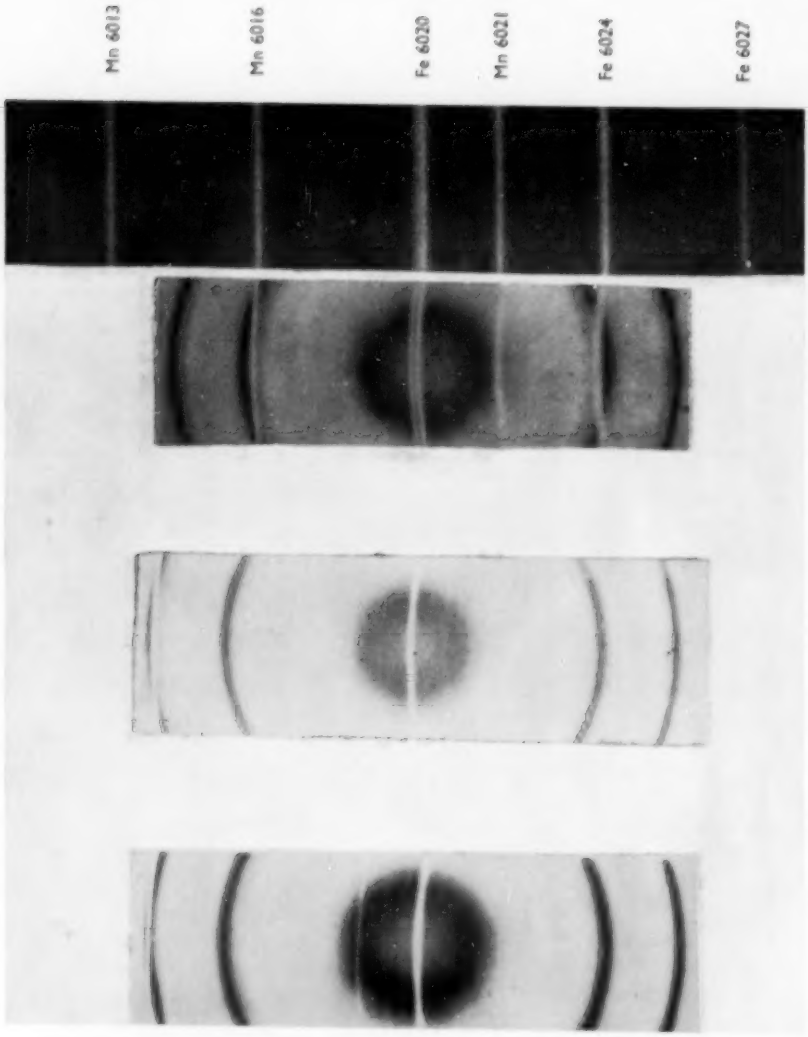
Summary

Abt has discussed the possible influence of hyperfine structure on solar line profiles. Allowing for hyperfine structure, and with the aid of detailed assumptions concerning physical conditions on the Sun, he reproduced the Utrecht profile of the weak manganese line λ 5516·8. This paper describes the interferometric investigation of the moderately strong manganese lines $\lambda\lambda$ 6021, 6016, 6013, and the iron line λ 6027. The method of circular channels was used, but with the line under examination at the centre of the heterochromatic fringe system. A solar absorption coefficient is deduced from the iron line profile, and this is used to compute the profiles of the manganese lines allowing for hyperfine structure. The only major assumption necessary concerning solar conditions is that the lines are all formed in the same level of the solar atmosphere. The computed and observed profiles agree within the limits of experimental error, and are substantially different from the manganese line profiles computed on the assumption of no hyperfine structure.

Introduction.—Hyperfine structure in atomic spectra is attributed to the interaction of the orbital electrons of the atom with nuclear magnetic and electric moments. The largest effects are observed when the nucleus possesses a large magnetic dipole moment. Abt (1) has discussed the possible influence of hyperfine structure on certain solar lines of manganese, vanadium and some other elements. He found some correlation between the half-widths of the solar profiles and the total hyperfine splitting of the line as observed in the laboratory. Abt also found evidence that hyperfine structure influenced the curve of growth. More particularly, he computed the profile of the line Mn 5516·8 allowing for hyperfine structure (known from laboratory experiments) and assuming an excitation temperature, a kinetic temperature, a turbulence velocity and a damping constant for the solar atmosphere. He used a Milne–Eddington model of the solar atmosphere with pure scattering. After allowing for the instrumental pattern, he found it possible to reproduce the Utrecht profile of the line. A very large number of assumptions about the solar atmosphere are involved in this computation, and further, the line λ 5516·8 is rather weak (Rowland intensity 0).

A more convincing approach seems to be to compare profiles broadened by hyperfine structure with iron line profiles of similar intensity. No hyperfine structure has ever been detected in the iron spectrum.

This paper describes the investigation of the profiles of three moderately strong manganese lines and an iron line with a view to establishing the influence of hyperfine structure on the manganese lines. Manganese possesses only one stable isotope whose nucleus has a spin of $\frac{5}{2}$ and a magnetic dipole moment of 3·47 nuclear magnetons, so that large hyperfine structures ($\sim$0·1 A) may be expected.



(*a*) *Normal spectrum.*
(*b*) *Interferogram with Fe* 6020 *central (illustrating gain in resolving power).*
(*c*) *Interferogram with Fe* 6027 *central.*
(*d*) *Interferogram with Mn* 6013 *central.*

1. *Experimental method*

1.1. *The choice of lines.*—The manganese lines chosen should possess large hyperfine structures which are known from laboratory observations or which may be reliably calculated. They should be reasonably intense, to allow an accurate enough determination of the profiles, and for the same reason they should be in a clear region of the spectrum. This latter consideration eliminates most of the violet manganese lines. It seems unlikely that the observed doubling at Mn 4030·8, 4033·1, 4034·5 is due to hyperfine structure, for the observed separation is about twice the total hyperfine splitting measured in the laboratory. The most satisfactory lines seemed to be the triplet

$$3d^5 4s(^7S)4p^6P_{3/2,\,5/2,\,7/2} - 3d^5 4s(^7S)5s^6S_{5/2},\ \lambda\lambda\, 6013{\cdot}5,\ 6016{\cdot}6,\ 6021{\cdot}8,$$

of Rowland intensity 6.

The iron lines should be of similar intensity to the manganese lines and in the same spectral region. They should be of pressure class *a* or *b* so that, presumably, they do not suffer from excessive collision broadening. The only suitable iron line was Fe 6027·1, Rowland intensity 4. This line had the additional advantage that the excitation potential of the lower term is about the same as that of the manganese lines.

1.2. *The observational method.*—The Oxford solar spectroscope was used in conjunction with a Fabry–Perot etalon to achieve the necessary resolving power. The method of circular channels due to Treanor (**5**) was adopted, but with a new relative position of the channels and the line under observation. The axis of the etalon was at such an angle to the axis of the light beam emerging from the spectroscope that the line being examined fell at the centre of the heterochromatic fringe system. This centre was made bright by small adjustments to the angle of the etalon. For photometric studies this method avoids the difficulty of photometering along a narrow circular channel. Further, the absorption line is long enough to allow a long photometer slit, which helps to reduce the effect of plate grain. This arrangement of the method of circular channels for line profile studies was suggested by Dr M. G. Adam. Treanor used a field lens in the focal plan of the spectroscope. This was not employed, as the field of view under examination was always quite small, and the removal of the field lens eliminated an appreciable quantity of scattered light. The focal lengths of the second collimator and camera lenses were 15 cm and 47 cm respectively. The etalon plates were mounted with their wedge angles parallel to the spectroscope slit, and a small enough slit height was used so that the ghost images were completely clear of the central part of the primary fringe system. The focal plane of the camera had an aperture 5 cm long by 0·8 cm high, so that four solar exposures could be taken on a single 2 in. × 4 in. plate, as well as the intensity marks. Plate 4 shows the appearance of the photographs.

1·5 mm and 2·5 mm spacers were used and the reflectivity of the silver films on the etalon plates was about 93 per cent. This gave theoretical resolving powers of $1{\cdot}8 \times 10^5$ and $3{\cdot}0 \times 10^5$ for the two spacers respectively, while photographs of the cadmium line λ6438 indicated actual resolving powers of about $1{\cdot}2 \times 10^5$ and $2{\cdot}0 \times 10^5$. The difference between theoretical and actual resolving powers is a result of plate imperfections and possibly the limitation of the number of reflected beams by the etalon aperture.

For the 1·5 mm spacer the spectroscope slit width was 0·25 mm and for the 2·5 mm spacer, 0·10 mm. These were the largest slit widths allowable for the order-to-order scattered light to be kept sufficiently low. Ilford Rapid Process Panchromatic plates were used, and the exposure times were 2·5 min for the 1·5 mm spacer and 8 min for the 2·5 mm spacer. All exposures were of the centre of the solar disk.

The plates were calibrated by taking a photograph with the etalon removed and a calibrated step wedge mounted immediately in front of the spectroscope slit. A similar photograph without the step wedge was taken to make possible a correction for the non-uniform illumination of the slit. The exposure times for all photographs on the same plate were the same, a condition achieved by the use of filters and the adjustment of the spectroscope slit width for the calibration photographs.

The plates were stored for 24 h before processing, to allow the initial fading of the latent image to take place. They were developed in Parkhurst developer.

1.3. *Photometry.*—Microphotometer (**4b**) tracings were made across the centre of the heterochromatic fringe system at right angles to the absorption line with both low (6·89×) and high (48·2×) magnifications. The purpose of the former was to locate the precise position of the absorption line with respect to the centre of the heterochromatic system so that the dispersion could be calculated (see below). The width of the image of the photometer slit on the plate was 0·057 mm and the height 0·4 mm.

The characteristic curve of each plate was determined by photometering along each step of the step-wedge photograph parallel to the dispersion and through absorption lines. By making use of points in the lines a large number of points on the characteristic curve were obtained. The correction for non-uniform illumination of the spectroscope slit was derived from a tracing of the photograph without the step wedge, at right angles to the dispersion. By means of the characteristic curves thus obtained the photometer tracings of the lines were reduced to observed profiles.

2. *Treatment of the observed profiles*

2.1. *Determination of the dispersion.*—At the centre of the heterochromatic fringe system the dispersion in a monochromatic fringe is equal to the dispersion in the spectrum obtained in the focal plane of the camera lens without the etalon. The dispersion in a monochromatic interference fringe is given by the expression

$$d\lambda/dx = \lambda x/f^2,$$

where f mm is the focal length of the camera lens, and x mm the distance from the centre of the system. Now the distance between heterochromatic and monochromatic centres is

$$p = f^2 w/\lambda$$

(see Treanor (**5**)), where w A/mm is the dispersion without the etalon. Thus when $x = p$, $w = d\lambda/dx$. If, however, $x = p + z$, where z is the distance between the heterochromatic centre and the line under consideration, the dispersion in the line is

$$d\lambda/dx = (p + z)\lambda/f^2 = w(1 + z/p).$$

p was calculated to be 81·8 tracing squares with the large magnification, while z varied from 0 to 7 tracing squares. w was measured on tracings of a spectrum without etalon and found to be 0·096 A/tracing square on the high-magnification

tracings, 0·673 A/square on low-magnification tracings, and 0·58 A/mm on the plate. z was measured on the low-magnification tracings and $d\lambda/dx$ calculated for both low and high magnifications. The observational profiles were thus determined with the correct dispersion.

2.2. *The scattered light correction.*—The general scattered light in the spectroscope is known to be about 1·5 per cent with a slit height of 19 mm. With the present slit height of 3 mm the general scattered light was therefore negligible.

In the method of circular channels there is always a certain amount of light scattered from one heterochromatic order to another because the slit image of the spectroscope extends over more than one order of interference. This occurs even with quite narrow slits, since the spectroscope is by no means a perfect image-forming device, and has a rather wide apparatus function. In order to estimate the proportion of scattered light, the intensity half-way between heterochromatic orders was measured. There are two contributions to this intensity. First, even if the spectroscope were a perfect image-former, there would be an inter-order intensity due to the interferometer alone. Secondly, there is the light that is due to the imperfection of the spectroscope. It is only the latter with which we are concerned, for the former is not scattered from order to order. The purely interferometric inter-order intensity was calculated, assuming the Airy formula for the intensity distribution, but remembering the finite slit width, and subtracted from the total to leave the true scattered light. So long as the slit width is small compared with the distance between orders, the true scattered light intensity falls off inversely as the square of the fractional order distance from the neighbouring order (see Bray (**2**)). Thus the scattered light at the line being measured is $\frac{1}{4}$ of the inter-order scattered light arising from the spectroscope imperfections. The measured values were:—

Inter-order intensity	1·5 mm spacer	11·1 per cent
	2·5 mm spacer	19·7
Pure interferometric (calculated)	1·5 mm spacer	0·9
	2·5 mm spacer	2·6
Inter-order true scattered light	1·5 mm spacer	10·1
	2·5 mm spacer	17·1
Scattered light at measured line	1·5 mm spacer	2·25
	2·5 mm spacer	4·2

The profiles corrected for scattered light yielded the values for central intensities and equivalent widths shown in Table I.

The agreement of the equivalent widths for the two spacers is evidence that the scattered light correction has at least the correct ratio for the two spacers. The disagreement of the central intensities suggests that the difference in resolving powers with the two spacers is significant for the profiles concerned.

2.3. *Averaging the results from the two spacers.*—It is desirable to average the results from the two spacers. To achieve this, each point on each 1·5 mm profile was multiplied by a factor representing the difference in resolving power between the two spacers. This factor was not difficult to compute, as all the profiles were found to be quite closely Gaussian. The function turning a 1·5 mm profile into a 2·5 mm profile must, therefore, also be Gaussian. The function was found from the Fe 6027 and Mn 6021 profiles, for which the best data were available, and applied to the others. The average profile for each line was then

calculated. Fig. 1 shows the observed points obtained for Fe 6027, together with the mean profile. The final mean central intensities and equivalent widths of the profiles, with the root mean square errors of the mean, were as follows:

	Central intensities	Equivalent widths (A)
Fe 6027	$0{\cdot}468 \pm 0{\cdot}012$	$0{\cdot}065 \pm 0{\cdot}0013$
Mn 6021	$0{\cdot}412 \pm 0{\cdot}012$	$0{\cdot}089 \pm 0{\cdot}0015$
Mn 6016	$0{\cdot}437 \pm 0{\cdot}022$	$0{\cdot}086 \pm 0{\cdot}0031$
Mn 6013	$0{\cdot}479 \pm 0{\cdot}012$	$0{\cdot}083 \pm 0{\cdot}0018$

TABLE I

Line	Central intensities		Equivalent widths	
	1·5 mm	2·5 mm	1·5 mm	2·5 mm
Fe 6027	0·494	0·511	0·065	0·059
	0·408	0·430	0·067	0·068
	0·533	0·433	0·058	0·064
	0·502		0·065	
	0·530		0·068	
Mean	0·493	0·458	0·065	0·064
Mn 6021	0·470	0·463	0·088	0·081
	0·430	0·424	0·091	0·092
	0·433	0·370	0·091	0·088
Mean	0·444	0·418	0·090	0·088
Mn 6016	0·447		0·082	
	0·507	0·374	0·082	0·097
	0·517		0·084	
Mean	0·490	0·374	0·083	0·097
Mn 6013	0·512		0·085	
	0·531	0·480	0·081	0·082
	0·502		0·085	
	0·551		0·074	
Mean	0·524	0·480	0·081	0·082

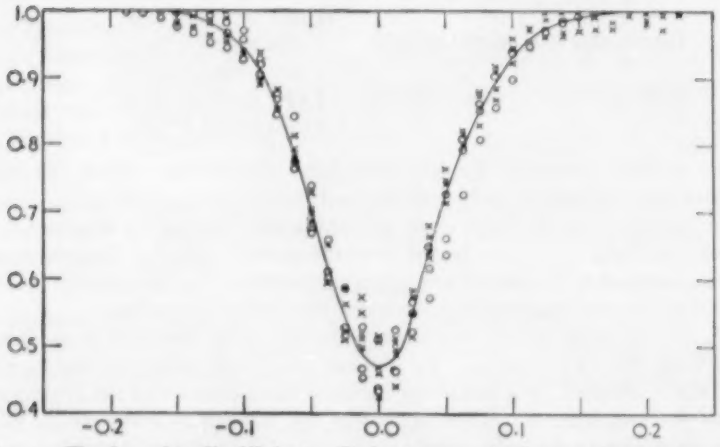


FIG. 1.—*The observed profile of Fe* 6027. *Circles represent points from* 2·5 *mm spacer exposures, and crosses points from* 1·5 *mm spacer exposures, converted as described in the text, Section* 2.3. *Ordinates are intensities as a fraction of the continuum, and abscissae are wave-length differences in angstroms.*

2.4. *Systematic errors.*—There may well be a number of systematic errors in the profiles. The methods of computing the scattered light and of averaging the results from the two spacers may not be thoroughly sound. However, these errors are likely to be the same for all four lines, and since the object is a comparison of the profiles and not an absolute determination, systematic errors are unlikely to be of much importance.

3. *Computation of manganese profiles from the iron profile*

3.1. *Deduction of the solar absorption coefficient from the Fe profile.*—The Milne–Eddington model with pure scattering was selected. For this model the residual intensity r for observations at the centre of the solar disk is given by

$$r = \frac{1 + cn^{-1} + \frac{1}{2}\sqrt{3}(\frac{2}{3}cn^{-1} - 1)(\xi + \frac{1}{2}\sqrt{3})^{-1}(1 - \xi^2)(1 + \xi\sqrt{3})^{-1}}{1 + c}, \qquad (1)$$

where $n = 1 + s_v/k_0$, $\xi = 1/n$, k_0 = mass coefficient of continuous absorption, s_v = mass scattering coefficient, and c = coefficient of limb darkening. From the data of Moll, Burger and van der Bildt, $c = 0{\cdot}962$. To solve the equation for n, the method used by Plaskett (**4a**) for a more complicated problem was adapted. The equation may be written in the following form:

$$\frac{1}{n} = \frac{r(1 + c) - \{1 - \frac{1}{2}\sqrt{3}(1 - \xi^2)(\xi + \frac{1}{2}\sqrt{3})^{-1}(1 + \xi\sqrt{3})^{-1}\}}{c + \frac{1}{2}\sqrt{3} \cdot \frac{2}{3}c(1 - \xi^2)(\xi + \frac{1}{2}\sqrt{3})^{-1}(1 + \xi\sqrt{3})^{-1}}.$$

The right-hand side of this equation was evaluated for $\xi = 0{\cdot}0$, $0{\cdot}2$, $0{\cdot}4, \ldots 1{\cdot}0$ for values of r in the observed profile in the iron line. For each value of r a graph of ξ against $1/n$ was plotted. The value of ξ for which $\xi = 1/n$ was then selected, and this value of $1/n$ was the solution of the equation. $n - 1$, which is proportional to the solar scattering coefficient, was thus determined as a function of wave-length. The curve possessed a very slight asymmetry, which may have been accidental. The two halves were averaged to produce a symmetrical scattering coefficient.

3.2. *The hyperfine structure of the manganese lines.*—The manganese lines λλ 6013, 6016, 6021 have not been thoroughly investigated in the laboratory, but the hyperfine structures of the terms from which they arise have been deduced by White and Ritschl (**6**) from measurements on other lines. The structure of the upper term has also been determined by Fisher and Peck (**3**). The application of selection and intensity rules then permits the evaluation of the structures of the lines λλ 6013, 6016, 6021. There is no evidence of perturbations in any of the laboratory investigations, so that there is no reason to doubt that the intensity rules are obeyed. The hyperfine structures thus evaluated are given in Table II and shown diagrammatically in Fig. 2.

3.3. *Computation of the manganese profiles.*—The absorption coefficient of each hyperfine structure component was given the same shape as that of the iron line but a magnitude proportional to the relative intensity of the component. In this way an absorption coefficient for the whole line was evaluated. The absolute magnitude of the coefficient was adjusted so that when the coefficient was substituted into equation (1) the equivalent widths of the computed and observed profiles agreed. After two or three attempts a computed profile of the correct equivalent width was produced. In Fig. 3 the observed and computed profiles are displayed. The slight dip on the long-wave side of Mn 6021 is probably the predicted line Fe 6021·82; this dip was not included in the equivalent width for normalization purposes. It is seen that the agreement between the observed and computed profiles is well within the random experimental error.

TABLE II

Mn 6021		Mn 6016		Mn 6013	
Wave-length difference (mA)	Relative intensity	Wave-length difference (mA)	Relative intensity	Wave-length difference (mA)	Relative intensity
0·0	0·7	0·0	16·7	0·0	100·0
29·0	12·1	25·5	24·2	19·0	61·4
29·5	1·6	33·5	100·0	36·0	32·7
52·5	18·1	47·5	24·5	47·5	12·7
54·0	2·2	52·5	54·5	52·5	20·5
64·0	100·0	62·0	19·4	61·0	28·6
70·5	1·8	67·5	25·5	64·0	27·3
71·5	19·2	75·5	9·1	68·5	19·1
81·5	72·5	76·5	10·6	79·0	9·1
82·0	16·5	83·0	1·8	85·0	8·2
91·5	11·0	88·0	16·7	89·0	5·5
94·5	50·4	93·5	10·6	94·5	2·3
99·5	33·0	94·5	24·2		
102·0	10·3	95·5	24·5		
103·0	19·8	96·5	19·4		

FIG. 2.—*The hyperfine structure pattern of the manganese lines. The length of each line represents the relative intensity of the component.*

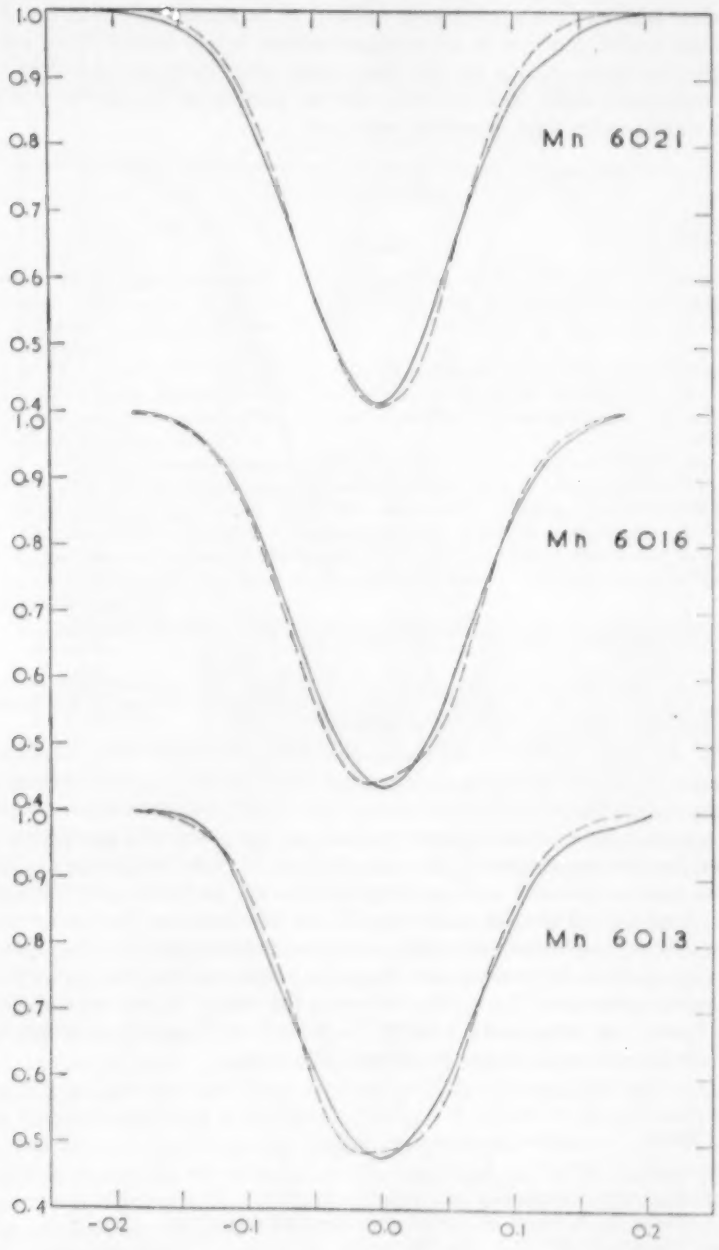


FIG. 3.—*The profiles of the manganese lines. The continuous lines are the observed profiles, and the broken lines the computed profiles. Ordinates are intensities as a fraction of the continuum, and abscissae are wave-length differences in angstroms.*

Since the three lines have significantly different central intensities but similar equivalent widths, such agreement is unlikely to be accidental. Further, Fig. 4 shows the profile that one of the manganese lines would have if there were no hyperfine structure—this is the iron line profile with the same equivalent width as the manganese lines—and it is clear that the manganese line profiles are being substantially modified by hyperfine structure.

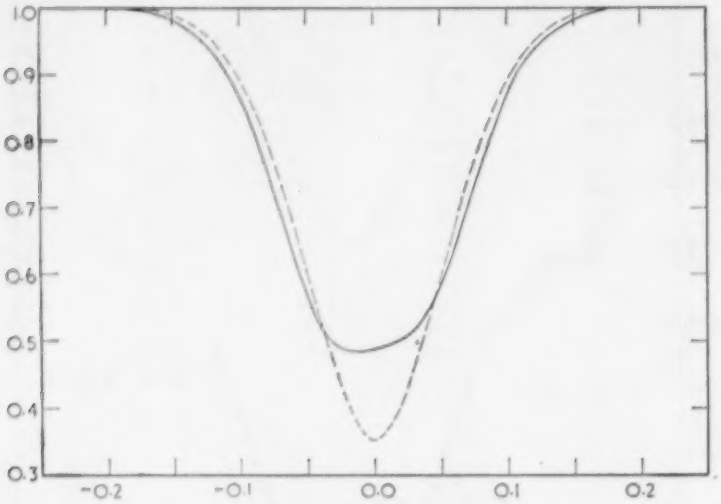


FIG. 4.—*The profile of Mn* 6013 *computed with hyperfine structure (continuous line) and without hyperfine structure (broken line). Ordinates are intensities as a fraction of the continuum, and abscissae are wave-length differences in angstroms.*

The observed profiles of the manganese lines have thus been satisfactorily accounted for by the inclusion of hyperfine structure as a cause of broadening. It is emphasized that the only major assumption lying behind the above procedure for comparing the iron and manganese line profiles is that the manganese lines and the iron line are formed in the same level of the solar atmosphere. This is likely to be approximately true, as the intensities and excitation potentials of the lines are similar. The solar model is no doubt deficient, and for this reason the solar scattering coefficient deduced for the iron line probably has little meaning. The comparison of the profiles must, however, be insensitive to the model chosen.

Acknowledgments.—The method of using the centre of the heterochromatic fringe system was suggested by Dr M. G. Adam, and I am very grateful to her for many discussions during the course of this work.

University Observatory,
Oxford:
1955 *March* 16

References

(1) Abt, A., *Ap. J.*, **115**, 199, 1952.
(2) Bray, R. J., *M.N.*, **114**, 540, 1954.
(3) Fisher, R. A. and Peck, E. R., *Phys. Rev.*, **55**, 270, 1939.
(4a) Plaskett, H. H., *M.N.*, **91**, 870, 1931.
(4b) Plaskett, H. H., *M.N.*, **95**, 160, 1934; *Comm. Univ. Obs., Oxford*, No. 4.
(5) Treanor, P. J., *M.N.*, **109**, 389, 1949; *Comm. Univ. Obs., Oxford*, No. 28.
(6) White, H. E. and Ritschl, R., *Phys. Rev.*, **35**, 1146, 1930.

CONTINUUM INTENSITIES IN PLANETARY NEBULAE

M. J. Seaton

(Received 1955 January 22)*

Summary

Thirty-four planetaries are considered for which continuum intensities have been measured, T_e may be calculated from the [O III] intensities and N_e may be calculated from forbidden line intensities or from the surface brightness.

Values of $B_c/H\delta$ measured by Aller (B_c = Balmer continuum intensity) are in reasonable agreement with theory. About half of the nebulae considered appear to come closer to Case A than to Case B (A = optically thin, B = optically thick in Lyman lines).

Total continuum intensities are calculated for H II regions allowing for: (1) all H and He II recombination and free–free continua and (2) the H $2s \rightarrow 1s$ two-quantum continuum, collisional deactivation of H 2s being taken into account. The calculated intensities, relative to Hδ, are compared with the measurements of Page for both sides of the Balmer limit. The mean results are in fair accord but there is a considerable scatter probably due to observational errors.

Recently measured Balmer discontinuities (Δ) are in good agreement with theory for 6 nebulae, but for NGC 6826 Δ (obs) is significantly smaller than Δ (calc). This object is unusual in that it is surrounded by a faint outer envelope. The small value of Δ (obs) is probably due to the conversion of Ly α into H $2s \rightarrow 1s$ quanta in an extensive H I region. When such processes are not important Δ (obs) may be used to improve the estimates of T_e and N_e.

---

1. *Introduction.*—Spectra of planetary nebulae reveal the presence of bright emission lines and of background continua. The present paper attempts a detailed comparison between theory and observation for the continuum intensities. This provides a useful check on our understanding of the physical processes and on our knowledge of the physical conditions in these objects.

The most complete published observations are those of Page (**1, 2**), who studied the region from 5000 to 3000 A. The continuum observed for $\lambda > 3647$ A was termed by Page the "visual continuum", V_c. For $\lambda < 3647$ A the intensity is much greater due to the Balmer continuum of hydrogen. From the discontinuity at 3647 A we may estimate the intensity per angstrom unit of the Balmer continuum at the Balmer threshold, which in the present paper will be denoted† by B_c. Page reported large systematic differences between theory and observation for the ratio $B_c/H\delta$. If confirmed, such differences would imply that the theory used to calculated the hydrogen line intensities was incorrect. The question is reconsidered in the present paper using measurements of Aller (**3, 4, 5**) for $B_c/H\delta$.

* Received in original form 1954 April 27.

† The intensity of the Balmer continuum is frequently denoted by Ba_c but Page (**2**) denotes by Ba_c the *total* intensity for $\lambda < 3647$ A.

The best test for the theory of the continuum intensities is provided by recent measurements of the Balmer discontinuities made by Minkowski (**6**), by Barbier and Andrillat (**7**) and by Aller (**8**). An apparent discrepancy between the measured discontinuity for NGC 7027 and that to be expected using the electron temperature calculated from the relative intensities of [O III] lines has been pointed out by Aller (**8**). The question is discussed in the present paper using improved [O III] intensities and allowing for deactivation effects.

The processes most likely to be of importance for the continua are (i) recombination and free–free transitions and (ii) two-quantum emission from certain metastable states.

The H-recombination continuum*, for which the emissivity may be calculated accurately, may be expected to be the most important process of type (i). After hydrogen the next most abundant element is He; according to Wurm (**9**) the He/H abundance ratio is about $\frac{1}{8}$. Recombination to highly excited states of He will contribute to the continuum in the observable spectral region. Exact calculation of this contribution would be difficult, but since the states concerned will be nearly hydrogenic we may estimate that the total intensity of the He I-recombination continuum will not exceed $\frac{1}{4}$ of the H-recombination continuum. This contribution will be neglected. In certain high-excitation objects the He II-recombination continuum may be more important due to the greater attractive power of the He^{++} field. In the present paper we consider the H-recombination and the He II-recombination continua and neglect all other recombination continua. The intensity of the H^--attachment continuum has been estimated by Greenstein and Page (**10**), who consider it unlikely to be of importance.

The continuum resulting from 2s→1s transitions in atomic hydrogen by two-quantum emission, which has been considered by Kipper (**11**) and by Spitzer and Greenstein (**12**), may be expected to be the most important process of type (ii). This will be referred to as the 2q-continuum. Possible contributions from similar transitions in He and in He^+ will be neglected. The semi-classical calculations of Purcell (**13**) for collisional deactivation of the 2s metastable state of hydrogen, together with recent estimates of the densities in planetary nebulae, suggest that the 2q-emissivity may depend critically on the electron density. A detailed quantum-mechanical study of the deactivation process has therefore been made (**14**), and the results obtained, which should be of a sufficient accuracy, are employed in the present paper.

2. *Summary of theory formulae*

2.1. *Hydrogen line intensities.*—Bohm and Aller (**15**) have shown that the electron velocity distribution in planetary nebulae will be very close to Maxwellian. For a given electron temperature T_e the hydrogenic continuum emissivities are readily calculated, but certain additional assumptions have to be made in order to calculate the line emissivities. It is to be expected that conditions will always lie somewhere between the two extreme cases considered by Baker and Menzel (**16**); Case A refers to a nebula optically thin in the Lyman lines and Case B to a nebula optically thick in the Lyman lines. For both A and B it is assumed that the stellar Lyman line radiation is absorbed in the stellar atmosphere and lower nebular layers.

* The total intensity resulting from recombination of electrons to X^+ ions, together with free–free transitions of electrons in the fields of X^+ ions, will be referred to as the X-recombination continuum.

The number of bare nuclei with charge Z_i (H^+, He^{++}, etc.) per cm^3 is denoted by N_i and the number of electrons by N_e. The energy emitted in transitions $n \to n'$ by the corresponding hydrogenic ions (H, He^+, etc.) is then given by

$$I_{nn'} = 1{\cdot}42 \times 10^{-22} \frac{g_{nn'}}{(nn')^3} Z_i^4 \frac{b_n \exp\{X_n Z_i^2\}}{t_e^{3/2}} N_i N_e \text{ erg cm}^{-3} \text{ sec}^{-1}, \quad (1)$$

where $t_e = 10^{-4} T_e$ and $X_n = 15{\cdot}78/n^2 t_e$. Baker and Menzel (**16**) give the coefficients b_n as functions of T_e (Cases A and B) and also the Kramers–Gaunt factors $g_{nn'}$.

In the present paper all calculations will be made in the first place assuming Case B for both H and He II. Conversion factors for various combinations of Cases A and B will then be given in Section 2.8.

2.2. *The He^{++} abundance.*—Denoting the α-particle density by N_α and the proton density by N_p, Wurm (**9**) gives*

$$N_\alpha / N_p = \Pi / 17{\cdot}2\xi, \quad (2)$$

where

$$\Pi = I(\text{He II}, \lambda 4686)/I(\text{H}\beta, \lambda 4861) \quad (3)$$

and

$$\xi = e^{3X_4} b_4(\text{He II})/b_4(\text{H}). \quad (4)$$

The quantity ξ has been calculated by Wurm with the following results (Case B):

$t_e = 10^{-4} T_e$	ξ
0·5	0·37
1·0	0·40
2·0	0·44

It will be shown in Section 3 that t_e lies between 1·0 and 2·0 in most planetaries. We therefore take $\xi = 0{\cdot}42$, giving

$$N_\alpha / N_p = 0{\cdot}14\Pi. \quad (5)$$

The hydrogen Balmer lines $n \to 2$ are blended with the He II lines $2n \to 4$. It is readily shown that this does not have any marked effect on the Balmer decrements.

2.3. *Hydrogenic recombination and free–free continua.*—The emission intensities are calculated in a form which is most readily compared with the observations of Page (**1**, **2**); denoting the emission intensities per angstrom unit by I_λ we put

$$J_\lambda = 10^3 I_\lambda / \text{H}\delta. \quad (6)$$

The contribution to J_λ arising from recombination to state n of a hydrogenic ion is given by

$$J_n^{(Z_i)} = 9{\cdot}22 g_{n\kappa}^{(Z_i)} \frac{Z_i^4}{n^3} e^{X_n Z_i^2} \frac{N_i}{N_p} \Psi(\nu, t_e), \quad (7)$$

where

$$\Psi(\nu, t_e) = \frac{\nu^2}{b_6} \exp\{-(0{\cdot}438 + 1{\cdot}439\nu)/t_e\} \quad (8)$$

and ν is the wave number in reciprocal microns (i.e. $\nu = 10^4/\lambda$ with λ in A). For the Kramers–Gaunt factors $g_{n\kappa}^{(Z_i)}$ we take $g_{2\kappa}^{(1)} = 0{\cdot}872$, $g_{4\kappa}^{(2)} = 0{\cdot}925$ and put all other $g_{n\kappa}^{(Z_i)}$ equal to unity. The error arising from this approximation will be of order 5 per cent or less (Menzel and Pekeris (**17**)).

* The author is indebted to Dr K. Wurm for informing him of this work in advance of publication.

The contribution to J_λ arising from free–free transitions in the fields of ions of charge Z_i is

$$J_{ff}^{(Z_i)} = 0{\cdot}292\bar{g}Z_i^2 t_e \frac{N_i}{N_p}\Psi(\nu, t_e) \tag{9}$$

with (17)

$$\bar{g} = 1 + 0{\cdot}078\left(\frac{\nu}{Z_i}\right)^{1/3}(1 + 1{\cdot}39t_e/\nu). \tag{10}$$

We consider the wave-length range corresponding to $2{\cdot}0 \leqslant \nu \leqslant 3{\cdot}0$; the Balmer limit is at $\nu = 2{\cdot}74$. Within this range the H-recombination contribution to J_λ is

$$J_\lambda(\mathrm{H}) = \sum_{n=n_0}^{\infty} J_n^{(1)} + J_{ff}^{(1)} \quad \begin{pmatrix} n_0 = 3, & \nu < 2{\cdot}74 \\ n_0 = 2, & \nu > 2{\cdot}74 \end{pmatrix} \tag{11}$$

and the He II-recombination contribution is

$$J_\lambda(\mathrm{He\ II}) = \sum_{n=n_0}^{\infty} J_n^{(2)} + J_{ff}^{(2)} \quad \begin{pmatrix} n_0 = 5, & \nu < 2{\cdot}74 \\ n_0 = 4, & \nu > 2{\cdot}74 \end{pmatrix}. \tag{12}$$

The functions $J_\lambda'(\mathrm{H})$ ($J_\lambda(\mathrm{H})$ for Case B) and $J_\lambda'(\mathrm{He\ II})$ ($J_\lambda(\mathrm{He\ II})$ for Case B and $N_\alpha/N_p = \frac{1}{4}$) are given in Tables I and II. In preparing these tables the summations over n for $n \geqslant 10$ were replaced by integrals which are readily evaluated.

TABLE I

		J_λ' (H)				
λ	ν \ t_e	0·5	1·0	1·5	2·0	3·0
5000 A	$2{\cdot}0\,\mu^{-1}$	$0{\cdot}5_2$	1·5	2·3	3·0	4·0
4545	2·2	$0{\cdot}3_6$	1·4	2·3	3·2	4·9
4167	2·4	$0{\cdot}2_4$	1·2	2·3	3·3	5·2
3846	2·6	$0{\cdot}1_6$	1·1	2·2	3·3	5·6
3647^+	$2{\cdot}74^-$	$0{\cdot}1_2$	1·0	2·1	3·3	5·7
3647^-	$2{\cdot}74^+$	24·1	17·5	15·9	15·2	15·7
3571	2·8	21·2	16·7	15·7	15·2	15·9
3333	3·0	13·8	14·5	14·9	15·2	16·6

TABLE II

		J_λ' (He II)				
ν	t_e \	0·5	1·0	1·5	2·0	3·0
$2{\cdot}0\,\mu^{-1}$		2·2	2·9	3·6	4·1	5·1
2·2		1·5	2·7	3·6	4·3	5·6
2·4		1·0	2·4	3·5	4·4	6·1
2·6		$0{\cdot}6_5$	2·1	3·4	4·5	6·5
$2{\cdot}74^-$		0·4	1·9	3·3	4·5	6·8
$2{\cdot}74^+$		13·1	10·6	10·5	10·8	12·0
2·8		11·5	10·2	10·4	10·8	12·2
3·0		7·5	8·8	9·9	10·7	12·7

2.4. *The H 2s two-quantum continuum.*—Spitzer and Greenstein (12) give the emissivity of the 2q-continuum proportional to a quantity X which is less than or equal to unity. In calculating X they considered the following two processes by which H atoms could enter the 2s state:

Process α.—Atoms enter the 2s state by capture or by cascade from higher discrete states. X is then defined by

$$X = \frac{\text{Number of 2s}\rightarrow\text{1s transitions by the 2q process}}{\text{Number of recombinations on all levels with } n \geqslant 2}. \qquad (13)$$

Neglecting deactivation, Spitzer and Greenstein estimated X to be between 0·30 and 0·35 and adopted the value 0·32. Allowing for deactivation X is then given by

$$X = \frac{0{\cdot}32}{1 + WN_e/A_{2q}}, \qquad (14)$$

where WN_e is the probability per second of collisionally induced 2s→2p transitions and A_{2q} the total radiative transition probability for the 2q process. According to Spitzer and Greenstein $A_{2q} = 8{\cdot}227\ \text{sec}^{-1}$. Using the results of Breit and Teller (**18**) for electron impact producing 2s→2p transitions, Spitzer and Greenstein concluded that WN_e/A_{2q} was small compared to unity. However, Purcell (**13**) has recently shown, using semi-classical methods, that collisions with protons will be more important, a result which is confirmed by quantum-mechanical calculations (**14**). Assuming equal densities of electrons and protons we obtain*

$$W = 5{\cdot}3 \times 10^{-4}\ \text{cm}^3\ \text{sec}^{-1} \quad \text{for} \quad T_e = 1{\cdot}0 \times 10^4\ \text{deg. K}$$

and

$$W = 4{\cdot}7 \times 10^{-4}\ \text{cm}^3\ \text{sec}^{-1} \quad \text{for} \quad T_e = 2 \times 10^4\ \text{deg. K}.$$

Consideration of the probable sources of error in the quantal method used leads to the conclusion that the results for W should be correct to within 20 per cent. Similar calculations have been carried out for 3s→3p transitions in Na produced by electron impact, for which good agreement with experiment was obtained.

Adopting the mean value $W = 5{\cdot}0 \times 10^{-4}\ \text{cm}^3\ \text{sec}^{-1}$ we obtain finally

$$X = \frac{0{\cdot}32}{1 + 0{\cdot}6 \times 10^{-4} N_e}. \qquad (15)$$

Process β.—Atoms may enter the 2s state by collisionally induced 2p→2s transitions. For this process to be important a sufficiently large number of atoms must enter the 2p state by absorption of Ly α quanta from the ground state. The parameter X is now defined by

$$\frac{X}{1-X} = \frac{\text{Total number of 2s}\rightarrow\text{1s transitions by the 2q process}}{\text{Total number of Ly } \alpha \text{ quanta escaping}}. \qquad (16)$$

The value of X depends on the total amount of neutral hydrogen, and in consequence a precise calculation becomes much more difficult; we may therefore consider some limiting conditions. If Case A is applicable, then process β is certainly unimportant. Spitzer and Greenstein concluded more generally that process β was negligible for processes occurring in regions where most of the hydrogen was ionized. Allowing for the larger value of W adopted in the present paper it would still appear that this conclusion is justified. On the other hand, if the ionized part of the nebula is surrounded by a sufficiently extensive neutral hydrogen region, X defined by (16) may approach the value unity.

It is by no means certain that most planetaries are surrounded by H I regions (Wurm and Singer (**19**)), and for a number of objects there are some grounds for

* These results are about one-half of the published results of Purcell. The difference is largely due to the fact that Purcell did not allow for the reduced mass.

believing that Case A may be applicable (Wurm (**9**), and Section 4 below). For a first survey of the problem of the sources of the continuous emissions we will therefore neglect process β; its importance may be reconsidered if other processes prove inadequate to account for the observed continuum intensities.

The 2q contribution to $J_\lambda, J_\lambda'(2q)$, is given in Table III calculated for Case B with $X = 0{\cdot}32$.

TABLE III

ν	t_e	0·5	1·0	1·5	2·0	3·0
		J_λ' (2q)				
$2{\cdot}0\,\mu^{-1}$		0·9	$1{\cdot}0_5$	1·1	1·2	1·3
2·2		1·3	1·5	1·6	1·7	1·8
2·4		1·8	2·0	2·1	2·3	2·4
2·6		2·3	2·6	2·8	3·0	3·1
2·74		2·8	3·1	3·4	3·5	3·7
2·8		3·0	3·3	3·6	3·8	4·0
3·0		3·7	4·1	4·5	4·7	5·0

2.5. *Corrections for space reddening.*—The calculated relative intensities of the Balmer lines are comparatively insensitive to T_e and to the particular model assumed. To a first approximation it may be assumed that the true relative Balmer emission intensities are the same for all planetary nebulae and that variations in the observed relative intensities are due to differences in the amount of space reddening. Berman (**20**) obtained a correction for space reddening by plotting the observed Balmer intensities against an effective path-length through the obscuring galactic stratum. In this way he obtained average decrements corrected for reddening which were found to be in good agreement with the theory.

Since there is considerable patchiness in the distribution of the obscuring clouds, we shall adopt the procedure of determining the amount of reddening from a comparison of theory and observation for the Balmer decrements.

In the wave-length range of interest the relation between reddening and absorption may be represented with sufficient accuracy by

$$\log\left(I_T(\lambda)/I_O(\lambda)\right) = c(\nu - \nu_0) \qquad (17)$$

(Whitford (**21**), Divan (**22**)), where $I_T(\lambda)$ is the true emission intensity, $I_O(\lambda)$ the observed intensity and c the reddening constant. We assume the optical properties of all obscuring clouds to be the same (Divan (**22**)) and take $\nu_0 = 0{\cdot}34$ (Whitford (**21**), Berger, Divan and Fringant (**23**)). Calculations will be made in the first place assuming Case B, for which the calculated value of c will be denoted by c'; if Case A is applicable $c = c' + 0{\cdot}21$, the additive constant being practically independent of T_e.

2.6. *Determination of electron temperatures.*—Electron temperatures may be calculated from the relative intensities of the [O III] lines (Menzel, Aller and Hebb (**24**), Seaton (**25**)). We obtain

$$t_e = 1{\cdot}43/\{\log\left(I(4959+5007)/I(4363)\right) - 0{\cdot}942 + \log\left(1 + 0{\cdot}0438 \times 10^{-4} N_e t_e^{-1/2}\right)\}, \qquad (18)$$

where $t_e = 10^{-4} T_e$ (T_e in deg. K). In using this formula we allow for absorption and calibration errors by taking $I(4959+5007)$ relative to $H\beta = 100$ and $I(4363)$ relative to $H\gamma = 50$ (**25**). The deactivation term involving N_e is in all cases small but is none the less significant.

2.7. *Determination of electron densities.*—Electron densities in planetary nebulae may be determined by two independent methods. The first utilizes the relative intensities of forbidden lines (Seaton (**25**)) and the second utilizes the absolute emission rate in the hydrogen spectrum as estimated from the absolute surface brightness and the nebular dimensions (Menzel and Aller (**3**)). The forbidden line method, which will be referred to as method (a), can at present only be applied to a limited number of objects, since in most cases insufficient observational material is available. For our present purpose the densities obtained by method (a) are to be preferred, since, if local condensations occur, this method gives an average local density effective for deactivation processes. It also has the additional advantage that it is not necessary to make any assumptions about the nebular distances or nebular geometry and that the results are practically independent of the theory assumed for the hydrogen lines.

Considerable improvement in the measurement of the surface brightnesses of nebulae have recently been made, using photoelectric techniques, by Liller and Aller (**26**) and* by Liller (see Aller (**8**)). However, the published densities tend to be too small due to neglect of space absorption. Denoting the observed average surface brightness of Hβ by S_β, and using equation (17), we obtain for the true surface brightness $10^{1.72c}S_\beta$. For a spherically symmetric nebula we may readily obtain from (1)

$$\log\left(\int N_e N_p A^2 dA\right) = \log S_\beta + 1{\cdot}72c + 2\log A_1 + 11{\cdot}47 + 2D - \log r, \qquad (19)$$

where S_β is in erg cm^{-2} sec^{-1}, A_1 is the angular radius in seconds of arc adopted for the measurement of S_β (the measured flux being proportional to $A_1^2 S_\beta$), r is the distance of the nebula in parsecs, N_e and N_p are the local electron and proton densities in cm^{-3}, which may be functions of A, the angular distance from the centre measured in seconds of arc and finally

$$D = \tfrac{1}{2}\log[t_e^{3/2}/b_4 e^{X_4}]. \qquad (20)$$

TABLE IV

t_e	D'
0·5	0·041
1·0	0·176
2·0	0·316
4·0	0·467

Table IV gives D as a function of T_e for Case B, denoted by D'; if Case A is applicable $D = D' + 0{\cdot}121$, the additive constant being practically independent of T_e.

In employing (19) we will put $N_e = N_p$, an approximation which will certainly be justified if the level of ionization is low. The most serious departures from this relation may be expected when nearly all the He is doubly ionized. We then have $N_e \simeq N_p + 2N_\alpha$ with $N_p \simeq 4N_\alpha$. However, in this case we should also allow for the He II contribution to Hβ; in (19) $N_e N_p$ should be replaced by $N_e(N_p + 0{\cdot}6N_\alpha)$. It readily follows that the values of N_e calculated assuming $N_e = N_p$ should be multiplied by 1·15 and hence that the error is small.

In employing this expression we follow Liller and Aller in assuming the density to be uniform in the spherical shell defined by $A_2 \leqslant A \leqslant A_1$ and to be zero

* The results obtained by Liller (**8**) have been adopted in preference to those of Liller and Aller (**26**) since they are apparently the more accurate. The differences between the results of these two determinations are, however, in all cases small.

outside this domain (note that A_2 may equal zero, in which case we have a uniform sphere). With this assumption we obtain

$$\log N_e = 5{\cdot}97 + \tfrac{1}{2}\log S_\beta - \tfrac{1}{2}\log(A_1 r) + 0{\cdot}86c + D - \tfrac{1}{2}\log[1-(A_2/A_1)^2]. \quad (21)$$

The use of this relation, together with the values of S_β, A_1, A_2 and r adopted by Liller and Aller or by Liller, will be referred to as method (b). It is readily shown that the densities calculated from (19) assuming Case A will be greater than those calculated assuming Case B by a factor of 2·0.

For nebulae for which there is insufficient observational material to use either method (a) or method (b) we may obtain approximate density estimates using the photographic magnitudes m_n. Denoting the photographic magnitude in monochromatic Hβ light by m_β we may put

$$m_\beta = m_n + \Delta m(\mathrm{H}\beta). \quad (22)$$

The quantity $\Delta m(\mathrm{H}\beta)$ may be expected to depend on the nebular spectrum and in particular to vary with the ratio $I(4959+5007)/\mathrm{H}\beta$; however, the values of $\Delta m(\mathrm{H}\beta)$ calculated by the method of Menzel and Aller (**3**) show this dependence to be comparatively small. Since $m_\beta = -2{\cdot}5\log(A_1^2 S_\beta) + \text{const.}$, we have

$$\log S_\beta + 2\log A_1 + 0{\cdot}4m_n = -0{\cdot}4\Delta m(\mathrm{H}\beta) + \text{constant}. \quad (23)$$

Using the results of Liller and Aller or of Liller for S_β together with the values of m_n quoted by Vorontsov-Velyaminov (**27**) it was verified that the left-hand side of (23) showed no systematic dependence on $\log(I(4959+5007)/\mathrm{H}\beta)$. The variation of $\Delta m(\mathrm{H}\beta)$ was therefore neglected. Adopting the mean for the 15 nebulae studied by Liller and Aller and by Liller we obtain

$$\log S_\beta + 2\log A_1 + 0{\cdot}4m_n = 4{\cdot}35 \pm 0{\cdot}31, \quad (24)$$

where $\pm 0{\cdot}31$ is the root mean square deviation. In view of the larger errors involved we adopt the simple approximation of considering the nebulae as uniform spheres with diameters in seconds $d = 2A_1$. This gives

$$\log N_e = 8{\cdot}60 - 0{\cdot}2m_n - \tfrac{3}{2}\log d - \tfrac{1}{2}\log r + D + 0{\cdot}86c. \quad (25)$$

The use of (25), together with the values of m_n, d and r listed by Vorontsov-Velyaminov (**27**) and assuming Case B, will be referred to as method (c). A similar equation quoted in a previous paper (**28**) was obtained using a conversion from m_n to S_β essentially similar to that of Menzel and Aller (**3**). This equation gives densities greater than those obtained from (25) by a factor of 2.

It may be noted that the errors involved in using (25) may in certain instances be considerably greater than are indicated by the r.m.s. deviation in (24). The most extreme case is NGC 2440, for which method (a) gives $N_e = 4\times10^4$ cm^{-3} and method (c) gives $N_e = 4\times10^2$ cm^{-3}. The relative exposures (RE) of Curtis (**29**) give a rough measure of the surface brightness. A plot of $\log(\mathrm{RE})$ against $m_n + 5\log d$ suggests that $m_n + 5\log d$ is greatly overestimated for this object. It also suggests that $m_n + 5\log d$ is somewhat overestimated for I 418, and this is confirmed by the photoelectric measures.

2.8. *Summary of theory relations.*—Fortunately the ratios $b_n(\mathrm{B})/b_n(\mathrm{A})$ are very insensitive to T_e, and in consequence conversion factors from Case B to Case A may be obtained which are practically independent of T_e. The final expression for the continuum intensities, $J_\lambda = 10^3 I_\lambda/\mathrm{H}\delta$, may be written in the form

$$J_\lambda = \sigma J_\lambda'(\mathrm{H}) + 0{\cdot}55\,\Pi\tau J_\lambda'(\mathrm{He\,II}) + \frac{\sigma}{1+0{\cdot}6\times10^{-4}N_e} J_\lambda'(2q), \quad (26)$$

where $J_\lambda'(\mathrm{H})$, $J_\lambda'(\mathrm{He\,II})$ and $J_\lambda'(2q)$ are given in Tables I, II and III and where σ and τ are given in Table V for various combinations of Cases A and B. According to Wurm (**9**) the combination of Case B for H and Case A for He II is very improbable.

TABLE V

Case, H	B	A	A	B
Case, He II	B	B	A	A
σ	1·00	1·46	1·46	1·00
τ	1·00	0·84	1·46	1·74

3. *The observational material*

3.1. *The data of Table VI.*—We consider 34 planetaries for which continuum intensity measurements have been made and for which estimates of T_e and N_e may be obtained. Table VI gives the following data:

(i) The catalogue numbers (right ascensions and declinations are given by Aller (**5**) and by Vorontsov-Velyaminov (**27**)).

(ii) The space reddening constants c' calculated from a comparison of the observed relative intensities of $H\beta$, $H\gamma$, $H\delta$, H_9 and H_{10} with the theoretical relative intensities for Case B with $t_e = 1{\cdot}5$. We do not use H_7 and H_8 since they are frequently blended with other features.

(iii) The ratio $\mathrm{II} = I(4686)/I(4861)$ (II may be assumed to be small when no value is given).

(iv) Values of $t_e = 10^{-4}T_e$ calculated from (18).

(v) Values of N_e calculated by methods (a), (b) or (c).

(vi) The Balmer continuum intensity relative to Hδ according to Aller (**3**, **4**, **5**). The observed ratio, $(10^3\mathrm{B_c/H\delta})_{\mathrm{obs}}$, must be corrected for reddening using (17) and for blending with He II using (1), (5) and (7). The quantity tabulated is

$$\log(10^3\mathrm{B_c/H\delta}) = \log(10^3\mathrm{B_c/H\delta})_{\mathrm{obs}} + 0{\cdot}30c' + \log\left(\frac{1+0{\cdot}08\,\mathrm{II}}{1+0{\cdot}29\,\mathrm{II}}\right). \qquad (27)$$

The last term on the right of (27) is small (less than 0·10 in all cases).

3.2. *Accuracy of the calculated electron densities.*—Comparison between the results obtained using methods (a) and (b) has been made elsewhere (**30**). It is seen that fair accord is obtained for NGC 6572 and for NGC 7027. For I 418 method (b) gives a density greater than that obtained from method (a) by a factor of at least 3·1; however, if in place of equation (21) we use equation (19) together with the density distribution function of Wilson and Aller (**31**), this factor is reduced to 1·6 (**30**). For NGC 7662 method (a) appears to give a density appreciably greater than that obtained from method (b), a result which might be explained in terms of local condensations. This will be discussed further in Section 6.3.

The r.m.s. deviation between the values of $\log N_e$ calculated from methods (b) and (c) is 0·2 but bigger errors may occur in certain cases. For one nebula of low density, NGC 40, methods (a) and (c) have been shown to be in fair accord (**28**).

4. *Balmer continuum intensities relative to Hδ.*—Equation (7) gives

$$\log(10^3\mathrm{B_c/H\delta}) = 0{\cdot}880 - (0{\cdot}190/t_e) - \log b_8. \qquad (28)$$

In Fig. 1 this function is plotted against T_e and compared with the observational results summarized in Table VI.

TABLE VI

Nebula, NGC, etc.	ϵ'	$\bar{\Pi}$	t_e	$10^{-4} N_e$	$\log\left(\frac{10^3 B_e}{H\delta}\right)$	Notes (See p. 289)
An. $0^h 25^m$	0·13	0·26	1·6	0·29 (c)	...	(1)
I 351	0·24	0·60	1·1	0·35 (c)	...	(1)
II 2003	0·11	0·56	1·2	0·4 (c)	...	(1)
1535	0·04	0·24	1·4	0·26 (b) B	1·62	(1, 2)
			1·4	0·5 (b) A		
I 418	0·24	...	1·8	0·8 (a)	1·21	(2, 3)
			1·75	2·5 (b) B		
			1·7	5 (b) A		
II 2165	0·19	0·60	1·5	0·28 (c)	1·12	(1)
J 900	0·21	0·47	1·3	0·17 (c)	0·82	(1)
2440	0·04	0·76	1·7	4 (a)	0·58	(1, 4)
II 3568	0·31	...	1·1	0·17 (c)	1·33	(1)
II 4593	0·08	...	0·8	0·34 (b) B	1·17	(1, 2)
			0·75	0·7 (b) A		
6210	0·64	...	1·0	3·1 (b) B	...	(2, 3)
			1·0	6 (b) A		
II 4634	0·39	...	1·1	0·4 (c)	1·18	(1)
6543	0·39	...	0·95	1·5 (b) B	1·42	(2, 3)
			0·9	3·0 (b) A		
6572	0·48	0·01	1·25	5 (a)	1·33	(2, 3, 5)
			1·3	1·9 (b) B		
			1·25	4 (b) A		
An. $18^h 47^m$	0·19	...	1·0	...	1·47	(1)
6741	0·26	0·34	1·3	0·5 (c)	...	(1)
II 4846	0·06	...	1·4	1·5 (c)	1·04	(1, 6)
6778	0·18	0·16	1·7	0·06 (c)	0·53	(1)
6790	0·61	...	1·3	10 (c)	1·42	(1, 6)
6803	0·05	0·08	1·2	0·5 (b) B	1·79	(1, 2)
			1·2	1·0 (b) A		
6807	0·15	...	1·4	0·8 (c)	...	(1)
6818	(−0·13)	0·82	1·55	0·25 (b) B	1·21	(1, 2, 7)
			1·5	0·5 (b) A		
6826	0·55	0·06	1·1	0·6 (b) B	1·48	(2, 3)
			1·1	1·2 (b) A		
6833	0·54	...	2·0	2·0 (c)	1·27	(1, 6)
6884	1·07	0·12	1·1	1·4 (c)	...	(1)
6886	0·66	0·40	1·3	0·9 (c)	...	(1)
7009	0·46	0·14	1·2	1·0 (b) B	1·31	(2, 3, 8)
			1·2	2·0 (b) A		
7026	0·46	0·28	0·8	0·19 (c)	1·17	(1)
7027	0·95	0·39	1·55	4 (a)	1·33	(2, 3, 9)
			1·55	4 (b) B		
			1·45	9 (b) A		
II 5117	0·41	0·14	1·4	1·9 (c)	...	(1, 6)
An. $21^h 31^m$	0·29	0·90	1·9	0·6 (c)	1·09	(1)
II 5217	0·21	0·13	1·1	1·1 (c)	1·13	(1)
An. $22^h 29^m$	(−0·11)	...	1·4	...	1·37	(1)
7662	0·20	0·53	1·3	5 (a)	1·11	(2, 3, 10)
			1·4	1·1 (b) B		
			1·4	2·2 (b) A		

Notes

(a) Densities obtained from relative intensities of forbidden lines (**25**).

(b) Densities obtained from photoelectric measurements of surface brightness (**26, 8**), Cases A and B being considered separately.

(c) Densities obtained from the surface brightness as estimated from the photographic magnitudes. Case B is assumed. The temperatures are not corrected for deactivation.

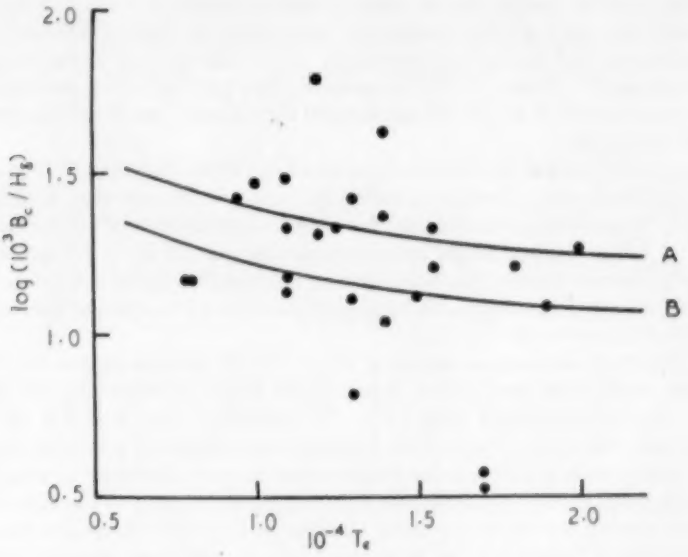


FIG. 1.—*log* ($10^3 B_c/H\beta$) *against* T_e. *Theory curves are given for Case A and Case B using* (28). *The observed values are plotted from Table VI.*

The main conclusion we draw from Fig. 1 is that there is no evidence for systematic differences between theory and observation. The present results differ from those of Page (**2**) because we have used: (i) higher electron temperatures (**25**); (ii) the Balmer intensities of Aller in place of those of Page (for a number of objects it appears that Page may have measured B_c too small; see Section 5 below); (iii) a different method of correcting for reddening (the method of Page would be expected to give a bigger scatter); (iv) a correction for blending with He II.

Of the 25 objects considered in Fig. 1, 11 come close to Case A, 9 come close to Case B and 5 are in poor agreement with theory. It is doubtful, however, if the observations are sufficiently accurate to decide which case is applicable for individual objects. The mean value of the quantity c' is 0·44 for the objects apparently coming close to Case A and is 0·24 for the objects apparently coming close to Case B. It would be somewhat surprising to find real differences in reddening between these two groups.* For Case A objects the correct space reddening constant is $c = c' + 0{\cdot}21$ (Section 2.5); we might therefore expect the

* The possibility cannot, however, be ruled out. Johnson (**32**) considers that there is evidence for a real correlation between certain physical and spatio-kinematic characteristics of planetary nebulae.

(1) Line intensities, Aller (**5**); (2) T_e from [O III] intensities adopted by Liller and Aller (**26**); (3) Line intensities, Aller (**4**); (4) T_e adopted by Seaton (**25**); (5) N_e calculated from the data of Liller and Aller (**8, 26**) differs from that given by Liller and Aller; (6) Stellar nebula with d uncertain ($d = 2''$ adopted); (7) T_e for E side, continuum for W side (**5**); (8) T_e calculated from the data of Liller and Aller (**26**) differs slightly from that given by Liller and Aller; (9) The value of c' is rather greater than that adopted by Aller (**37**), but Aller states that his value may be underestimated; (10) I (4959 + 5007)/Hβ from MacRae and Stock (**39**).

quantity c' to be *smaller* for the objects coming closer to Case A. It appears probable that some objects apparently come close to Case A due to c' being overestimated, and that others apparently come close to Case B due to c' being underestimated. However, if it be assumed that there are no systematic errors in the observations, it would still appear that Case A and Case B nebulae are about equally numerous.

For Cases A and B the calculated values of $\log(10^2 B_c/H\delta)$ differ by 0·17 and the calculated reddening corrections differ by 0·06. It is seen that it should, in principle, be possible to discriminate between Case A and Case B objects using this ratio. A better method might be to measure the ratio of B_c to the intensities of the higher Balmer lines, as has been done by Greenstein (**33**) for the Orion nebula. The difference between $\log(10^2 B_c/H_{15})$ for Cases A and B is 0·10 and the reddening correction is very small.

5. *The total continuum intensities of Page.*—In the present section we compare the total continuum intensities* measured by Page (**2**) relative to Hδ with the theory intensities obtained from (26). We consider Cases A and B for H but Case B only for He II. Page gives the maximum measured intensity shortward of the Balmer limit and the maximum measured intensity in the visual continuum. The maximum beyond the limit would be expected to occur at the limit and we therefore correct for reddening using the Balmer limit wave-length and compare with the theory results for the short-wave side of the limit, denoted by J_{3647-}. Page states that V_c (per A) is approximately constant in the range $3647 \leqslant \lambda \leqslant 5000$ A. Tables I, II, III and VI show this to be in fair accord with theoretical prediction. We compare the maximum value of $10^3 V_c/H\delta$ as given by Page with the values calculated for $\lambda = 4167$ A ($\nu = 2{\cdot}4$).

We define the Balmer discontinuity as

$$\Delta = \log[I_{3647-}/I_{3647+}], \tag{29}$$

where I_{3647+} is the total continuum intensity on the long-wave side of the Balmer limit. Since V_c is approximately constant we have

$$\Delta \simeq \log[I_{3647-}/V_c]. \tag{30}$$

The detailed results are given in Table VII. Tables I and V show that, considering H-recombination alone, J_{3647-} should always be greater than 15 for both Case A and Case B and for all values of T_e. Of the 21 measured values given in Table VII, 11 give $J_{3647-} \leqslant 15$. It is also seen that several of the nebulae with low measured values of J_{3647-} have measured values of Δ close to zero. These results are difficult to reconcile with the theory. However, a comparison of Aller's results for $10^3 B_c/H\delta$ with Page's results for $10^3(I_{3647-} - V_c)/H\delta$ suggests that the observational error in Page's results is rather large and that this probably accounts for the apparent discrepancy with the theory. The fact that the mean observed value of J_{3647-} comes nearer to the theory value for Case B than to that for Case A is probably without significance.

The comparison between theory and observation for V_c suggests that the processes considered are adequate to account for the main features of the observed visual continua but in view of the rather large scatter a more detailed discussion scarcely seems justified.

* Aller (**3**, **4**, **5**) does not give the total intensities.

TABLE VII

Nebula, NGC, etc.	J_{8647} Theory, Case B H	He II	2q	Total	Theory, Case A H	He II	2q	Total	Observed (Page)	J_{4167} Theory, Case B H	He II	2q	Total	Theory, Case A H	He II	2q	Total	Observed $10^3 V_c/H\delta$ (Page)
$0^h\ 25^m$	15·7	1·5	2·9	20	22·9	1·3	3·7	28	13	2·5	0·5	1·8	4·8	3·7	0·4	2·3	6·4	13
I 351	17·1	3·5	2·6	23	25·0	2·9	3·3	31	15	1·4	0·9	1·6	3·9	2·1	0·8	2·1	5·0	16
II 2003	16·7	3·3	2·6	23	24·4	2·8	3·1	30	11	1·6	0·9	1·6	4·1	2·3	0·8	2·0	5·1	8
6210	17·5	0·0	1·1	19	25·6	0·0	1·0	27	28	1·2	0·0	0·7	1·9	1·8	0·0	0·6	2·4	(1·5)
6543	17·9	0·0	1·6	19·5	26·6	0·0	1·6	28	26*	1·1	0·0	1·1	2·2	1·4	0·0	1·1	2·5	4·1*
6572	16·6	0·0	0·8	17	24·2	0·0	1·2	25	14	1·7	0·0	0·5	2·2	2·5	0·0	0·7	3·2	...
6741	16·4	2·0	2·5	21	24·0	1·7	3·0	29	10	1·8	0·6	1·6	4·0	2·6	0·5	1·9	5·0	8
II 4846	16·2	0·0	1·8	18	23·6	0·0	1·8	25	20	2·0	0·0	1·1	3·1	2·9	0·0	1·1	4·0	10
6790	16·4	0·0	0·5	17	24·0	0·0	3·6	28	28*	1·8	0·0	0·3	2·1	2·6	0·0	0·2	2·8	3·5*
6807	16·2	0·0	2·3	18·5	23·6	0·0	2·5	26	14	2·0	0·0	1·4	3·4	2·9	0·0	1·6	4·5	4
6818	15·8	4·7	3·0	23·5	23·2	3·9	3·8	31	13*	2·4	1·6	1·4	5·4	3·4	1·3	2·4	7·1	4·9*
6826	17·1	0·3	2·3	20	25·0	0·3	2·6	28	27*	1·4	0·1	1·5	3·0	2·0	0·1	1·7	3·8	4·2*
6833	15·2	0·0	1·6	17	22·1	0·0	1·5	24	40*	3·3	0·0	1·0	4·3	4·8	0·0	1·0	5·8	11·5*
6884	17·1	0·7	1·8	20	25·0	0·6	1·7	27	25	1·4	0·2	1·1	2·7	2·0	0·2	1·1	3·3	2·5
6886	16·4	2·3	2·2	21	24·0	1·9	2·3	28	16	1·8	0·7	1·4	3·9	2·6	0·6	1·5	4·7	2
7009	16·8	0·8	2·0	20	24·5	0·7	2·1	27	23*	1·6	0·2	1·2	3·0	2·3	0·2	1·3	3·8	4·1*
7027	15·8	2·2	1·0	19	23·1	1·8	1·4	26	6·9*	2·4	0·8	0·6	3·8	3·5	0·6	0·9	5·0	1·7*
II 5117	16·2	0·8	1·6	19	23·6	0·7	1·5	26	50	2·0	0·3	1·0	3·3	2·9	0·3	0·9	4·1	...
$21^h\ 31^m$	15·3	5·3	2·6	23	22·3	4·4	3·0	30	13	3·1	2·1	1·7	6·9	4·5	1·8	2·0	8·3	6
II 5217	17·1	0·7	1·9	20	25·0	0·6	2·0	28	13*	1·4	0·2	1·2	2·8	2·0	0·2	1·3	3·5	5·2*
7662	16·4	3·0	0·8	20	23·9	2·5	1·2	28	11*	1·8	0·9	0·5	3·2	2·6	0·8	0·8	4·2	3·5*
Mean				20				28	20, 21†				3·5				4·5	5·7, 4·7†

* Photometric measurements (given weight 1 in averaging). The remainder of the observations are visual estimates (given weight ½ in averaging).
† Averages of photometric measurements.

6. *The Balmer discontinuity*

6.1. *Theory relations.*—The theory expression for the Balmer discontinuity may be written

$$\Delta = \log \left\{ \frac{i_-(\mathrm{H}) + 0{\cdot}55\,\mathrm{II} i_-(\mathrm{He\,II}) + (X/0{\cdot}32) i_{0\cdot32}(2q)}{i_+(\mathrm{H}) + 0{\cdot}55\,\mathrm{II} i_+(\mathrm{He\,II}) + (X/0{\cdot}32) i_{0\cdot32}(2q)} \right\}, \qquad (31)$$

where i_- denotes an intensity on the short-wave side and i_+ on the long-wave side of the Balmer limit. The value of X is discussed in Section 2.4 and the other functions occurring in (31) are given in Table VIII. The factor $0{\cdot}55\mathrm{II}$ is obtained assuming that the same Case (A or B) applies for both H and He II. If other combinations apply this factor should be multiplied by τ/σ (Table V).

6.2. *Comparison of calculated and observed Balmer discontinuities.*—Observed and calculated Balmer discontinuities are compared† in Table IX. The results for NGC 6543, 6572, 6826, 7009 and 7662 are those obtained by Barbier and Andrillat (**7**), who state that the internal precision for a single spectrum is of order 0·05; the number of spectra obtained are given in parentheses. The result quoted for I 418 is that obtained by Minkowski (**6**) and the result for NGC 7027 is deduced from the statement of Aller (**8**) that the Balmer discontinuity is consistent with the theory of Spitzer and Greenstein for $T_e = 10^4$ deg. K.

TABLE VIII

t_e	i_-(H)	i_-(He II)	i_+(H)	i_+(He II)	$i_{0\cdot32}(2q)$
0·8	10·21	5·80	0·292	0·56	1·69
1·0	7·51	4·43	0·411	0·68	1·38
1·2	5·90	3·65	0·498	0·80	1·19
1·4	4·86	3·17	0·572	0·90	1·04
1·6	4·14	2·80	0·631	0·95	0·92
1·8	3·62	2·51	0·677	0·96	0·83
2·0	3·22	2·30	0·707	0·97	0·78

The units adopted are such that $10^{-25} \times i \times N_e N_p$ is equal to the emission rate in erg cm^{-3} sec^{-1} per reciprocal micron.

TABLE IX

NGC, etc.	Observed Δ	Calculated Δ (a)	(b) B	(b) A	$N_e \to 0$
I 418	0·48 ± 0·07	0·53	0·61	0·67	0·47
6543	0·98 (4)	...	0·89	1·02	0·72
6572	0·79 (1)	0·87	0·75	0·83	0·58
6826	0·61 (4)	...	0·74	0·79	0·66
7009	0·82 (2)‡	...	0·72	0·79	0·62
7027	0·70	0·68	0·68	0·76	0·50
7662	0·81 (2)	0·80	0·65	0·70	0·56
Mean	0·76	...	0·72	0·79	0·57$_8$

Note that NGC 6826 is excluded in calculating the means.

The Balmer discontinuities have been calculated assuming process β of Section 2.4 not to be operative, i.e. using (15) for the calculation of X. Calculations have been made using the densities obtained by method (a) and by method (b) assuming both Cases B and A. To illustrate the effect of the deactivation terms, calculations have also been carried out for the limiting case of $N_e \to 0$.

† A similar comparison has been made in a previous paper (**34**) using slightly less accurate values of T_e and N_e.

‡ Barbier and Andrillat state that NGC 7009 was observd low on the horizon; Dr D. Barbier informs the author that the observed value of Δ is of lower accuracy due to the large correction for differential atmospheric absorption.

It is seen that, with the exception of NGC 6826, the observations are in good agreement with the theory employed. It has previously been suggested (**34**) that the low value of Δ observed for NGC 6826 is due to the operation of process β. Long exposures obtained by Duncan (**35, 36**) show this nebula to be unusual in that it is surrounded by a faint and extensive outer envelope. It seems to be very probable that this outer envelope is an H I region shining by the light of the H 2s→1s two-quantum radiation. An attempt to confirm this conclusion by obtaining the spectrum of the outer envelope would appear to be well worth while.

6.3. *Use of the Balmer discontinuity to obtain more precise temperatures and densities.*—The intensity ratios of forbidden lines and also the observed surface brightness give relations between T_e and N_e. Provided that process β may be neglected, another such relation is provided by the measured Balmer discontinuity.

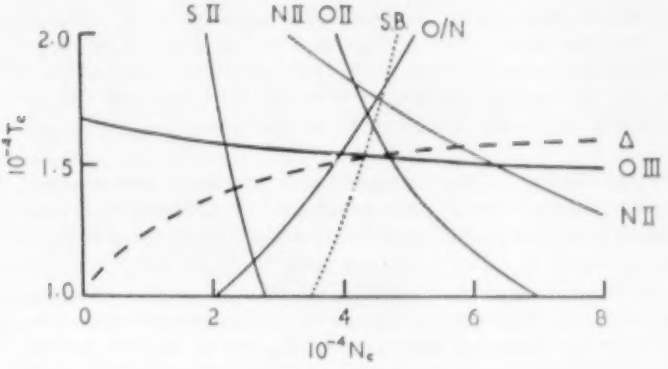


FIG. 2.—*T_e, N_e relations for NGC 7027.*

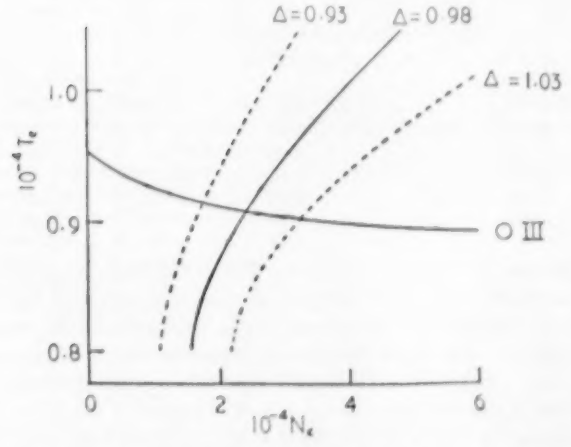


FIG. 3.—*T_e, N_e relations obtained from the [O III] ratio and from the Balmer discontinuity for NGC 6543. The observed Balmer discontinuity is $\Delta=0{\cdot}98$ with a probable error of about $\pm 0{\cdot}03$. To illustrate the sensitivity of the calculated density to errors in the measurement of Δ curves are given for $\Delta=0{\cdot}93$ and $\Delta=1{\cdot}03$.*

The relations between T_e and N_e for NGC 7027 are shown in Fig. 2. The curves marked S II, N II, O II and O/N are those given in a previous paper (**25**), the curve marked O III is that obtained from (18) together with the intensities adopted in Table VI, and that marked Δ is obtained from (31) with $\Delta = 0{\cdot}70$. We also give the relation obtained from the surface brightness, assuming Case B, which is marked S.B. It is seen that all these relations are in fair accord, although the S II curve indicates a density which is rather low. It is possible that a discrepancy for S II might arise from the fact that the atomic data are rather less accurate (**25**), but it seems more probable that the differences result from variations in the intensities across the disk. Aller (**37**) has recently given some intensities for a selected point in the brightest part of NGC 7027 which are somewhat different from those adopted for Fig. 2. Using these we obtain densities from S II which are appreciably *greater* than those obtained from O II or N II.

The Balmer discontinuity is particularly sensitive to N_e when T_e is small and N_e is not too large. This is illustrated in Fig. 3 for NGC 6543, for which apparently reliable values of T_e and of N_e may be obtained from the [O III] intensity ratio and the Balmer discontinuity. It should be noted that the errors in T_e due to errors in the [O III] ratio are much smaller when T_e is small than when it is large.

Improved estimates of the average local temperatures and densities, effective for excitation and deactivation processes, have been obtained from a combination of the data on forbidden line intensities and on the Balmer discontinuities. These are given in Table X. For I 418 and NGC 7662 the Balmer discontinuity is found to be in better agreement with the densities previously calculated from the forbidden lines than those calculated from the surface brightness, and in particular for NGC 7662 the measured discontinuity appears to confirm that the effective local density cannot be much less than that previously obtained. It appears that the average local density* exceeds, by a factor of at least 2, the root mean square density for the total volume obtained from the surface brightness.

TABLE X

NGC, etc.	$t_e = 10^{-4} T_e$	$10^{-4} N_e$
I 418	1·8	0·8
6543	0·91	2·4
6572	1·25	4
7009	1·21	2·8
7027	1·55	4·5
7662	1·30	5

6.4. *Variations of Δ across the disk.*—For NGC 7662 Page (**1**) obtained a distribution across the disk for V_e which was more uniform than that obtained for B_e. This is in satisfactory agreement with what would be expected from deactivation effects. For I 418 Minkowski (**6**) states that Δ decreases inside the bright ring. It is evident from Table IX that only a small decrease could result from a decrease in deactivation effects. A decrease in Δ might be explained in terms of an increase in T_e but from a study of the factors determining T_e Aller (**38**) has concluded that large variations in I 418 are improbable. It would be interesting to check this conclusion by measuring the variation of the [O III] ratio across the

* A typical average local density would be $\bar{N}_e = \int N_e^{S+1}\, dv / \int N_e^{S}\, dv$ with $S \geqslant 1$.

disk and at the same time to obtain more definite numerical results for the variation of Δ.

7. *Acknowledgments.*—The present paper was revised and new material added while the author was working at the Institut d'Astrophysique de Paris, on leave from the Department of Physics, University College, London.

Institut d'Astrophysique,
Paris :
1955 *January.*

References

(1) T. L. Page, *M.N.*, **96,** 604, 1936.
(2) T. L. Page, *Ap. J.*, **96,** 78, 1942.
(3) D. H. Menzel and L. H. Aller, *Ap. J.*, **93,** 195, 1941.
(4) L. H. Aller, *Ap. J.*, **93,** 236, 1941.
(5) L. H. Aller, *Ap. J.*, **113,** 125, 1951.
(6) R. Minkowski, *P.A.S.P.*, **65,** 161, 1953.
(7) D. Barbier and H. Andrillat, *Comptes Rendus*, **238,** 1099, 1954.
(8) L. H. Aller, *Astrophysics ; nuclear transformations, stellar interiors and nebulae*, pp. 212, 206, Ronald Press, New York, 1954.
(9) K. Wurm, *Mém. Soc. Roy. Sci. Liège*, **14,** 387, 1954.
(10) J. L. Greenstein and T. L. Page, *Ap. J.*, **114,** 106, 1951.
(11) A. Y. Kipper, *Tartu Astr. Obs. Publ.*, **32,** 63, 1952.
(12) L. Spitzer and J. L. Greenstein, *Ap. J.*, **114,** 407, 1951.
(13) E. M. Purcell, *Ap. J.*, **116,** 457, 1952.
(14) M. J. Seaton, *Proc. Phys. Soc.* A, **68,** 457, 1955.
(15) D. Bohm and L. H. Aller, *Ap. J.*, **105,** 131, 1947.
(16) J. G. Baker and D. H. Menzel, *Ap. J.*, **88,** 52, 1938.
(17) D. H. Menzel and C. L. Pekeris, *M.N.*, **96,** 77, 1936.
(18) G. Breit and E. Teller, *Ap. J.*, **91,** 215, 1940.
(19) K. Wurm and O. Singer, *Z. Astrophys.*, **30,** 387, 1952.
(20) L. Berman, *M.N.*, **96,** 890, 1936.
(21) A. E. Whitford, *Ap. J.*, **107,** 102, 1948.
(22) L. Divan, *Comptes Rendus*, **236,** 2301, 1953.
(23) J. Berger, L. Divan and A.-M. Fringant, *Comptes Rendus*, **236,** 2379, 1953.
(24) D. H. Menzel, L. H. Aller and M. H. Hebb, *Ap. J.*, **93,** 230, 1941.
(25) M. J. Seaton, *M.N.*, **114,** 154, 1954.
(26) W. Liller and L. H. Aller, *Ap. J.*, **120,** 48, 1954.
(27) B. A. Vorontsov-Velyaminov, *Gaseous nebulae and new stars*, pp. 564–575, Moscow, U.S.S.R. Academy of Sciences, 1948.
(28) M. J. Seaton, *Ann. d'Astr.*, **17,** 74, 1954.
(29) H. D. Curtis, *Publ. Lick Obs.*, **13,** 55, 1918.
(30) M. J. Seaton, *Comptes Rendus*, **239,** 1592, 1954.
(31) O. C. Wilson and L. H. Aller, *Ap. J.*, **114,** 421, 1951.
(32) H. M. Johnson, *Ap. J.*, **120,** 182, 1954.
(33) J. L. Greenstein, *Ap. J.*, **104,** 414, 1946.
(34) M. J. Seaton, *Ann. d'Astr.*, **17,** 296, 1954.
(35) J. C. Duncan, *P.A.S.P.*, **8,** 141, 241, 1936.
(36) J. C. Duncan, *Ap. J.*, **86,** 496, 1937.
(37) L. H. Aller, *Ap. J.*, **120,** 401, 1954.
(38) L. H. Aller, *Ap. J.*, **118,** 547, 1953.
(39) D. MacRae and J. Stock, *Nature, Lond.*, **173,** 589, 1954.

ON THE OCCURRENCE OF COMMENSURABLE MEAN MOTIONS IN THE SOLAR SYSTEM

II. The Mirror Theorem

A. E. Roy and M. W. Ovenden

(Communicated by the Director, University Observatory, Glasgow)

(Received 1954 December 6)

Summary

In Part I it is shown that the preference for near-commensurability of mean motions demonstrated in a previous paper extends to the small satellites of Jupiter, including the retrograde ones, and to the retrograde satellite of Saturn, implying that stability of near-commensurable configurations is the reason for such a preference. In Part II it is proved that if, at a certain epoch, a system of n gravitating point-masses has each radius vector from the (assumed stationary) centre of mass of the system perpendicular to every velocity vector (hereinafter called a "mirror configuration"), then the behaviour of each of the point-masses under the internal gravitational forces of the system after the epoch will be a mirror image of its behaviour prior to the epoch. It is further shown that, if a mirror configuration of the system exists at two separate epochs, then the orbit of each point-mass is periodic. The authors argue that such periodic orbits are the more stable (under the action of external forces) the shorter the interval of time between mirror configurations, and they show that the frequent occurrence of mirror configurations between any two point-masses requires that the mean motions of the two point-masses be nearly commensurable. In Part III, various near-commensurable pairs of orbits in the solar system are shown to behave according to the above arguments. The relationship of the present work to the more restricted "symmetry theorem" of Griffin is discussed.

Part I

1. *Introduction.*—In a previous paper (1), hereinafter referred to as Paper I, the present authors have shown that, in the solar system, among the planetary and satellite systems, the occurrence of commensurability between pairs of mean motions is more frequent than in a chance distribution. It was suggested that the observed distribution was due either to a property of the mechanism of formation of the solar system, or to an inherent stability of commensurable configurations.

More recent work has led to the adoption of the latter hypothesis. The authors were led to investigate this latter alternative by a consideration of bodies in the solar system not included in Paper I, namely Jupiter's satellites VI to XII and Saturn IX.

2. *The outer satellites of Jupiter and Saturn.*—Using the nomenclature of Paper I, we have

$$\epsilon = \left| \frac{n_2}{n_1} - \frac{A_2}{A_1} \right|, \qquad (2.01)$$

where n_1, n_2 are the mean motions of the two bodies, and A_1 and A_2 are integers (which, in Paper I, were less than or equal to 7).

The Sun disturbs Jupiter's satellites considerably, and the orbits of the retrograde satellites J VIII, J IX and J XI are not even approximately elliptic. Yet they have approximately equal mean motions, and if we consider the Sun as a satellite of Jupiter, the ratio of its mean motion to those of the retrograde ones is within the range of 1 to 7 previously considered. As is seen from Table I, these satellites are nearly commensurable in mean motion with the Sun, the ratio being 1 to 6. Indeed, the Sun's mean motion about Jupiter is almost exactly one-sixth of the mean of the mean motions of J VIII, J IX and J XI.

TABLE I

Satellites	$\frac{n_2}{n_1}$	A_2	A_1	$\frac{n_2}{n_1} - \frac{A_2}{A_1}$
Sun J VIII	0·170 55	1	6	+0·003 88
Sun J IX	0·171 96	1	6	+0·005 29
Sun J XI	0·159 82	1	6	−0·006 84
Sun and mean of J VIII, J IX, J XI	0·167 25	1	6	+0·000 59
Sun J VI	0·057 84	1	17	−0·000 98
Sun J VII	0·060 02	1	17	+0·001 20
Sun J X	0·058 67	1	17	−0·000 15
Sun and mean of J VI, J VII, J X	0·058 83	1	17	−0·000 01
Sun J XII	0·145 64	1	7	+0·002 78
Iapetus Phoebe	0·144 11	1	7	+0·001 25

Note: $\frac{1}{16} - \frac{1}{17} = 0{\cdot}003\,68$

$\frac{1}{6} - \frac{1}{7} = 0{\cdot}023\,81$

It is convenient to consider the other satellites of Jupiter here. The Sun's mean motion about Jupiter is nearly 1/7 of the mean motion of J XII, while the satellites J VI, J VII and J X form a group whose mean motions are nearly commensurable with the Sun's at 1/17, and once again their average mean motion is almost exactly commensurable with that of the Sun. Finally, Phoebe (Saturn IX), disturbed by Iapetus, pursues a retrograde orbit whose mean motion is commensurable with that satellite at 1/7. In Table I the mean motions are derived from the *Handbook of the British Astronomical Association for* 1953, except for those of J XII, Iapetus, Phoebe and the Sun. The last three were obtained from the *Connaissance des Temps* for 1954, while the mean motion of J XII was from Hilton and Brady's orbit No. IV (**2**).

These results provide additional evidence for the conclusions of Paper I, and suggest that the reason for the preference for near-commensurability is that such configurations are relatively stable. For, if the retrograde satellites of Jupiter and Saturn are captured bodies, it is difficult to understand why they should all have mean motions approximately commensurable with other bodies, unless they have reached positions of relative stability.

PART II

3. *The mirror theorem.*—In this section it is proved that, if n point-masses are acted upon by their mutual gravitational forces only, and at a certain epoch each radius vector from the (assumed stationary) centre of mass of the system is perpendicular to every velocity vector, then the orbit of each mass after that epoch is a mirror image of its orbit prior to that epoch.

Consider an isolated system of n point-masses, centre of mass G, and let their masses be m_i, $i = 1, 2, 3, \ldots n$. Suppose that at a certain time the radius vector from G to m_i be $\boldsymbol{R}_i$, etc., and the velocity vector of m_i be $d\boldsymbol{R}_i/dt$, etc. Then the equation of motion of m_i relative to G is given, in gravitational units, by

$$\frac{d^2\boldsymbol{R}_i}{dt^2} = -\sum_{j=1}^{n} m_j \frac{(\boldsymbol{R}_i - \boldsymbol{R}_j)}{[(\boldsymbol{R}_i - \boldsymbol{R}_j).(\boldsymbol{R}_i - \boldsymbol{R}_j)]^{3/2}}, \quad j \neq i. \tag{3.01}$$

The geometrical shapes and sizes of the orbits arising from this configuration, and their relative geometrical disposition, are completely specified by the masses m_i together with magnitudes and relative configuration of the radii vectores $\boldsymbol{R}_i$ and velocity vectors $d\boldsymbol{R}_i/dt$, i.e. by

$$\left.\begin{aligned} &m_i \text{ for all values of } i, \\ &\boldsymbol{R}_i.\boldsymbol{R}_j \text{ for all values of } i \text{ and } j, \\ &(d\boldsymbol{R}_i/dt).\boldsymbol{R}_j \text{ for all values of } i \text{ and } j, \\ \text{and} \quad &(d\boldsymbol{R}_i/dt).(d\boldsymbol{R}_j/dt) \text{ for all values of } i \text{ and } j. \end{aligned}\right\} \tag{3.02}$$

Consider now the geometrical shapes, sizes and relative geometrical disposition of the orbits arising from the same configuration of radii vectores and velocity vectors, but with the new independent variable $\tau = -t$. These orbits are completely specified by

$$\left.\begin{aligned} &m_i \text{ for all values of } i, \\ &\boldsymbol{R}_i.\boldsymbol{R}_j \text{ for all values of } i \text{ and } j, \\ &(d\boldsymbol{R}_i/d\tau).\boldsymbol{R}_j \text{ for all values of } i \text{ and } j, \\ \text{and} \quad &(d\boldsymbol{R}_i/d\tau).(d\boldsymbol{R}_j/d\tau) \text{ for all values of } i \text{ and } j; \end{aligned}\right\} \tag{3.03}$$

the equation of motion (3.01) being invariant under the transformation of t to τ. The geometrical shapes, sizes and relative disposition of the orbits arising from the configuration in τ are identical with those arising from the same configuration in t if conditions (3.03) are identical with conditions (3.02). Unless certain symmetry relationships hold between the masses m_i, conditions (3.03) are identical with conditions (3.02) if, and only if

$$(d\boldsymbol{R}_i/d\tau).(d\boldsymbol{R}_j/d\tau) = (d\boldsymbol{R}_i/dt).(d\boldsymbol{R}_j/dt) \text{ for all values of } i \text{ and } j, \tag{3.04}$$

and

$$(d\boldsymbol{R}_i/d\tau).\boldsymbol{R}_j = (d\boldsymbol{R}_i/dt).\boldsymbol{R}_j \quad \text{for all values of } i \text{ and } j. \tag{3.05}$$

But, from the definition of τ, (3.04) holds, and

$$(d\boldsymbol{R}_i/d\tau).\boldsymbol{R}_j = -(d\boldsymbol{R}_i/dt).\boldsymbol{R}_j \text{ for all values of } i \text{ and } j. \tag{3.06}$$

Condition (3.06) is equivalent to condition (3.05) if, and only if

$$(d\boldsymbol{R}_i/d\tau).\boldsymbol{R}_j = 0 = (d\boldsymbol{R}_i/dt).\boldsymbol{R}_j \text{ for all values of } i \text{ and } j. \tag{3.07}$$

Thus, if condition (3.07) holds at a certain epoch $t = \tau = 0$, the motion of the system of n bodies after this epoch is geometrically similar to the motion of the system prior to the epoch. At a mirror configuration the following conditions arise from the proof:—

(i) For all values of i, the geometrical shape and size of the orbit of the ith body after the mirror configuration is equivalent to the geometrical shape and size

of the orbit of the ith body prior to the mirror configuration, but described in the opposite sense.

(ii) For all values of i and j, the relative geometrical disposition of the ith and jth orbits after the mirror configuration is equivalent to the relative geometrical disposition of these orbits prior to the mirror configuration.

(iii) For all values of i, the ith orbit is continuous at the mirror configuration. These three conditions serve to define uniquely the nature of the "reflection" of the orbits.

4. *The two possible cases of a mirror configuration.*—Consider two of the n point-masses. Let NM_1A_1 be the plane containing G and the instantaneous velocity vector of m_1, and NM_2A_2 be the plane containing G and the instantaneous velocity vector of m_2, where M_1, M_2 are the projections, on the celestial sphere centred on G, of the radii vectores Gm_1 and Gm_2 respectively, and A_1, A_2 are the projections, on the same celestial sphere, of the instantaneous directions of motion of m_1 and m_2 respectively. Then, if $d\boldsymbol{R}_i/dt \,.\, \boldsymbol{R}_j = 0$,

$$M_2A_2 = M_2A_1 = M_1A_2 = M_1A_1 = 90^\circ.$$

Thus, either M_1 and M_2 are poles of the great circle A_1A_2, or A_1 and A_2 are poles of the great circle M_1M_2. Thus, if longitudes are measured along the plane of the appropriate orbit, from the common node N, for a mirror configuration either

(i) longitude of $M_1 = 0^\circ$ or 180°, longitude of $A_1 = \pm 90^\circ$ and longitude of $M_2 = 0^\circ$ or 180°, longitude of $A_2 = \pm 90^\circ$, or

(ii) longitude of $M_1 = \pm 90^\circ$, longitude of $A_1 = 0^\circ$ or 180° and longitude of $M_2 = \pm 90^\circ$, longitude of $A_2 = 0^\circ$ or 180°.

Since for any pair of point-masses, either (i) or (ii) must hold, for a mirror configuration of the n point-masses, it is necessary either (i) that all the radii vectores $\boldsymbol{R}_i$ are parallel or anti-parallel, and all the velocity vectors $d\boldsymbol{R}_i/dt$ are perpendicular to all the $\boldsymbol{R}_j$, or (ii) that all the velocity vectors $d\boldsymbol{R}_i/dt$ are parallel or anti-parallel, and all the radii vectores $\boldsymbol{R}_i$ are perpendicular to all the $d\boldsymbol{R}_i/dt$.

For any two of the n point-masses, the above cases correspond to a conjunction or opposition (with respect to G) with each body at either pericentron or apocentron. In case (i), the conjunction or opposition must occur at the common nodes of the two orbits, while in case (ii) the conjunction or opposition must occur 90° from the common node.

To distinguish more clearly between these two cases, we discuss the nature of the "reflection" of the orbits for each case separately.

The equation of motion (3.01) may be written

$$\left.\begin{aligned} \ddot{x}_i &= -\sum_{j=1}^{n} m_j \frac{(x_i - x_j)}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}}, \qquad j \neq i \\ \ddot{y}_i &= -\sum_{j=1}^{n} m_j \frac{(y_i - y_j)}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}}, \qquad j \neq i \\ \ddot{z}_i &= -\sum_{j=1}^{n} m_j \frac{(z_i - z_j)}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{3/2}}, \qquad j \neq i \end{aligned}\right\} \qquad (4.01)$$

where x, y, z are a rectangular set of axes with origin G, the centre of mass of the system of n point-masses.

Case (i):—Suppose that at $t = \tau = 0$, each m_i is on the line $y = 0$, $z = 0$, and is moving in a plane parallel to yGz. Suppose also that

$$(x_i,\ y_i,\ z_i,\ t) \qquad \text{etc.} \qquad (4.02)$$

is a solution of (4.01). Then

$$(\bar{x}_i = x_i;\ \bar{y}_i = -y_i;\ \bar{z}_i = -z_i;\ \tau = -t) \quad \text{etc.} \tag{4.03}$$

is also a solution of (4.01), which is unchanged by the substitution. Hence (4.02) and (4.03) represent possible orbits. The initial conditions of the set of orbits (4.02) are

$$(x_i, y_i, z_i;\ dx_i/dt, dy_i/dt, dz_i/dt;\ t) = (a_i, 0, 0;\ 0, g_i, h_i;\ 0) \quad \text{etc.} \tag{4.04}$$

By the transformation (4.03), these initial conditions for the set of orbits represented by (4.03) become

$$(\bar{x}_i, \bar{y}_i, \bar{z}_i;\ d\bar{x}_i/d\tau, d\bar{y}_i/d\tau, d\bar{z}_i/d\tau;\ \tau);\ = (a_i, 0, 0;\ 0, g_i, h_i;\ 0) \quad \text{etc.} \tag{4.05}$$

(4.04) and (4.05) are identical, and therefore (4.02) and (4.03) represent the same orbit. Hence the orbit of m_i for $t > 0$ is the mirror image of the orbit for $t < 0$ in the planes yGx and zGx.

Case (ii):—Suppose that at $t = \tau = 0$, each m_i is in the plane xGy and moving perpendicular to it. Suppose also that

$$(x_i, y_i, z_i, t) \quad \text{etc.} \tag{4.06}$$

is a solution of (4.01). Then

$$(\bar{x}_i = x_i;\ \bar{y}_i = y_i;\ \bar{z}_i = -z_i;\ \tau = -t) \quad \text{etc.} \tag{4.07}$$

is also a solution of (4.01), which is unchanged by the substitution. Hence both (4.06) and (4.07) represent possible orbits. The initial conditions of the set of orbits (4.06) are

$$(x_i, y_i, z_i;\ dx_i/dt, dy_i/dt, dz_i/dt;\ t) = (a_i, b_i, 0;\ 0, 0, f_i;\ 0) \quad \text{etc.} \tag{4.08}$$

By the transformation (4.07), these initial conditions for the set of orbits represented by (4.07) become

$$(\bar{x}_i, \bar{y}_i, \bar{z}_i;\ d\bar{x}_i/d\tau, d\bar{y}_i/d\tau, d\bar{z}_i/d\tau;\ \tau) = (a_i, b_i, 0;\ 0, 0, f_i;\ 0) \quad \text{etc.} \tag{4.09}$$

(4.08) and (4.09) are identical, and therefore represent the same orbit. Hence the orbit of m_i for $t > 0$ is the mirror image of its orbit for $t < 0$ in the plane xGy.

5. *The periodicity theorem.*—We define periodicity as follows:—In a system of n point-masses moving under their mutual gravitational forces only, their orbits are said to be periodic if, at periodic intervals of time, the same relative configuration of radii vectores and velocity vectors occurs with no change in scale.

Hence the periodicity theorem:—If n point-masses are moving under their mutual gravitational forces only, their orbits are periodic if, at two separate epochs, a mirror configuration occurs.

For, if the mirror condition (3.07) holds at $t = -t_0$ and at $t = 0$, then it holds at $t = +t_0$, the orbits after $t = 0$ being the mirror images of the orbits prior to $t = 0$. Hence the orbits are periodic, with period $2t_0$.

6. *Discussion of stability.*—The above section gives no information on the stability of such orbits against perturbations by external forces. The manner in which stability is treated here makes no pretence to being exhaustive or rigorous, but the present authors believe that it throws some light on the possible importance of the periodic orbits discussed above, and leads to conclusions which are subsequently confirmed by appeal to observation.

There are various accepted definitions of stability of an orbit. We will use Poincaré's conditions for stability (4), namely:

(i) the heliocentric distance of any planet cannot increase or decrease without limit,

(ii) the system repeatedly passes through the configuration it had at time t_0, say a times t_1, t_2, etc.,

(iii) close encounters of any pairs of planets are ruled out.

(The conditions for satellite orbits are analogous.)

Condition (iii) implies that a body occupying a stable orbit will not stray far radially from it.

The discussion in Section 5 does not of itself satisfy Poincaré's conditions. Although the orbits are periodic if the mirror condition is satisfied at two separate epochs, they might be unstable in the sense that a small change in their configuration would lead to non-periodic orbits and rapid departures from those orbits so far pursued.

Let us consider the classical case of a system consisting of two massive bodies of unequal masses M_1 and M_2 pursuing circular orbits about their common centre of gravity, with a small mass m moving about the larger massive body M_1, in the same plane as, and being perturbed by the other massive body, M_2. Initially, let M_1, M_2 and m satisfy the mirror condition, at time t_0 say. Then m will move in a perturbed orbit and as a consequence its line of apses will rotate. In general, one synodic period of M_2 and m later, the velocity vector of m will be no longer perpendicular to the radius vector of m. But if the initial distances of m and M_2 have been chosen correctly, M_2's perturbation of m may be such that the time of revolution of m from one pericentron to some future pericentron is equal to one or more synodic periods of M_2 and m. Then this new configuration, at t_1 say, satisfies the mirror condition again. Therefore the behaviour of m after t_1 is a mirror image in the conjunction line M_1mM_2 of its behaviour before t_1. Thus m will move in a disturbed periodic orbit, the perturbations having a period of $2(t_1-t_0)$. It seems likely that if mM_2 is such, initially, that the mirror configuration after t_0 does not arise to "reverse" the perturbations within a sufficiently short time, the disturbances might build up sufficiently to cause m to be captured by M_2, or to collide with it or M_1. Again, if the second mirror configuration does not arise sufficiently quickly after t_0, the strongly perturbed periodic orbit of m might take it close to M_2, when the perturbation of some fourth body might cause it to be captured by M_2 or sent to infinity.

The periodic orbits studied by G. H. Darwin (3) are of interest, and may be said to be examples of the mirror theorem. Darwin considers the classical coplanar three-body problem, and confines his attention to simple periodic orbits which are periodic after one synodic period only, although the possibility of more complex periodic orbits with periods which are multiples of the synodic period is mentioned. In addition to isolating certain periodic orbits which correspond to the orbits discussed in this section, Darwin examines the stability of such periodic orbits. His results (in particular those relating to his "Family A of planets") to some extent corroborate the suggestions made in this section.*

Hence the reasonable hypothesis is made that in this restricted three-body case, the periodic orbit of m will be the more stable against external perturbations the shorter the time-interval between the occurrence of mirror configurations.

7. *The effect of near-commensurability of mean motions of two bodies.*—One major effect of near-commensurability is the occurrence of large inequalities, since the amplitudes of long-period terms involve the factor

$$(A_2n_1-A_1n_2)^{-1}.$$

* We are indebted to the Referee for drawing our attention to this fact.

Let us consider, after Newcomb (**5**), the mean motions n_1 and n_2 of two bodies orbiting about a primary in the same plane, and so related that

$$A_2 n_1 - A_1 n_2 = 0, \tag{7.01}$$

A_1 and A_2 being integers, and $n_2 < n_1$. Then A_2 revolutions of the first body will require the same time as A_1 revolutions of the second body, and at the end of this period they will have returned to their original positions, having been in conjunction $(A_1 - A_2)$ times at the same number of equidistant points. Calling these points "conjunction points", we put

$$\nu = A_1 - A_2. \tag{7.02}$$

But if the commensurability (7.01) is not exact, so that

$$A_2 n_1 - A_1 n_2 = \kappa,$$

then the mean conjunction points will revolve about the origin with velocity k given by

$$k = \frac{A_2 n_1 - A_1 n_2}{\nu}, \tag{7.03}$$

k approaching zero as κ approaches zero.

In the cases which we shall discuss in detail in this paper, $A_1 = A_2 + 1$, and therefore $\nu = 1$, so that we are concerned with only one mean conjunction point revolving uniformly with velocity

$$k = A_2 n_1 - A_1 n_2. \tag{7.04}$$

By adjusting the mean motions, k can be given any value.

It was seen in Section 5 that, if the mirror condition is satisfied twice, the orbits are periodic, and the reasonable hypothesis was made that the periodic orbit of m would be the more stable the shorter the time-interval between the two satisfactions of the mirror condition. The shortest possible time between mirror configurations is obtained if one synodic period of m and M_2 equals the time taken by m's pericentron to move from one conjunction to the next. But, by (7.04), if the mean motions are not exactly commensurable, the line of conjunction revolves with a uniform velocity k, and a mirror configuration would occur at every conjunction if the mean motions of the bodies were such that

$$k = \dot{\omega},$$

where $\dot{\omega}$ is the rate at which m's pericentron revolves.

We can now relax the conditions of the classical system considered hitherto, and consider two bodies moving in non-coplanar, non-circular orbits about a primary, and perturbing each other. We discuss two major cases:—

(*a*) If the inclinations are small, so that the orbits are approximately coplanar, and one eccentricity is much larger than the other, a first approximation to the frequent occurrence of mirror configurations is for the mean motions of the two bodies to be so nearly commensurable that the rate of revolution of the line of apses of the orbit of greater eccentricity equals the rate of rotation of the conjunction line of the two bodies. A perfect mirror configuration would occur only when

$$\omega_1 = \omega_2, \quad \text{or} \quad \omega_1 = \omega_2 + 180^\circ,$$

where ω_1 and ω_2 are the longitudes of the pericentrons. Hence a libration will ensue, whose period is the synodic period of the pericentrons.

(*b*) If the inclinations are large, and the nodes of the orbits precess about the fundamental plane of the system at different rates, then by Section 4, a mirror configuration can occur only if conjunctions take place at either (i) the common

node of the two orbits, or (ii) 90° from the common node, with the bodies on the apse lines of their orbits in each case. In this case, frequent occurrences of mirror configurations can occur if the mean motions are such that conjunctions always take place according to (i) or (ii).

It is of interest to make an estimate of the relative importance of (*a*) and (*b*) in an actual case. If the two orbits are coplanar, and one is circular while the other has a small eccentricity e, then it can be shown that, if a conjunction (or opposition) occurs at a small angle ξ from the apse-line of the eccentric orbit, then the angle between the velocity vectors at conjunction (or opposition) is of the order of $e\xi$. On the other hand, if the two orbits are circular, and inclined at a small angle i to each other, then, if a conjunction (or opposition) (measured along one of the orbits) occurs at a small angle ξ from the common node of the two orbits (or at an angle $90° \pm \xi$ from the common node), then the angle between the radii vectores (or between the velocity vectors) is of the order of $\xi \sin i$. If the two orbits have small inclinations i_1 and i_2 ($i_2 < i_1$) to the fundamental plane of the system about which their nodes precess, then the mutual inclination of the two orbits will vary between the limits $(i_1 + i_2)$ and $(i_1 - i_2)$. Therefore, in a semi-quantitative way, we would expect that if

$$e > \sin i_1$$

case (*a*) would be the more important, while if

$$e < \sin i_1$$

case (*b*) would be the more important.

In Part III we inspect certain systems in the solar system with outstanding commensurabilities, and see how far they support the above arguments.

Part III

8. *Saturn's system*

(*a*) *Hyperion and Titan.*—From Table II of Paper I, $(n_H/n_T) = \frac{3}{4} - 0{\cdot}000\,57$. The orbits are very nearly coplanar, and the eccentricities of Titan and Hyperion are 0·029 and 0·104 respectively. Hyperion is very strongly disturbed by Titan. On the basis of the arguments of Part II, we would expect that the mean motions of the two bodies would be such that the rate of revolution of the conjunction point would be the same as that of Hyperion's perisaturnium. From (7.04), with $n_T = 22°{\cdot}577\,009$ per day (**7**), the point of conjunction retrogresses at a rate of 18°·66(9) per year. The rate of retrogression of the perisaturnium of Hyperion is 18°·66(3) per year (**7**). It is, indeed, well known (**9**) that Hyperion and Titan come into conjunction with each other only at or near the aposaturnium of the former. More accurate elements derived since Newcomb's time (**6, 7, 8**) confirm this result.

If Titan had a circular orbit, a mirror configuration would occur at every conjunction and opposition. However, due to the eccentricity of Titan's orbit, a perfect mirror configuration would occur only when Titan was also at its aposaturnium or perisaturnium. At other conjunctions, the satisfaction of the mirror condition is only approximate, and hence the reversal of perturbations at such conjunctions or oppositions is only approximate. A libration is therefore to be expected, whose period is the synodic period of the apse lines of Titan and Hyperion, i.e. 18·75 years (**7**), and a major libration of this period, of amplitude 14°·0 (**7**) is present in the motion of Hyperion. The system of Hyperion and Titan is thus in excellent agreement with the arguments of Part II.

(*b*) *Dione and Enceladus.*—The orbits are equally inclined to the fundamental plane (the plane of Saturn's equator), and retrogress at different rates. However, the inclinations are very small ($\sim 1'$). The eccentricities of Dione and Enceladus are 0·0022 and 0·0044 respectively. Enceladus is the less massive body, and suffers the greater perturbations.

For this system, a first approximation to the frequent occurrence of mirror configurations can be found by considering the orbits to be coplanar, and the orbit of Dione to be circular. Then, for a mirror configuration to occur at every conjunction, the rate of motion of the perisaturnium of Enceladus should equal the rate of motion of the point of conjunction. Now, from Table II of Paper I, $(n_D/n_E) = \frac{1}{2} + 0{\cdot}000\,643$. Therefore, from (7.04), with $n_E = 262^\circ{\cdot}731\,99$ per day (**7**), the point of conjunction advances by $123^\circ{\cdot}41$ per year. The rate of advance of the perisaturnium of Enceladus, $\dot{\omega}_E = 123^\circ{\cdot}43$ per year (**7**). Thus we have the relation noted by H. Struve (**10**) that

$$2n_D - n_E = \dot{\omega}_E$$

within the accuracy of the derived elements. Struve also noted that conjunctions of Enceladus and Dione occur always near the perisaturnium of Enceladus, about which point they oscillate.

A second approximation to the actual problem is to consider the orbits to be still coplanar, but to consider the eccentricity of Dione's orbit. Then mirror configurations at conjunction will be only approximate unless the conjunction occurs when Dione is also at its perisaturnium or aposaturnium. A libration is to be expected whose period is the synodic period of the perisaturnia of Enceladus and Dione. Now $\dot{\omega}_D = 30^\circ{\cdot}74$ per year (**6**), so that

$$4\dot{\omega}_D = \dot{\omega}_E$$

within the accuracy of the derived elements (**14**). We would thus expect a libration in the longitudes of Enceladus and Dione of period 3·88 years. According to the *Connaissance des Temps* (**7**), a libration of amplitude $11'{\cdot}24$ and period 3·89 years exists in the longitude of Enceladus, and a libration of amplitude $\sim 1'$ of the same period in the longitude of Dione.

A third approximation to the actual problem is to consider the inclinations of the orbits to the fundamental plane of the system. Exact mirror configurations can occur at conjunction only if the conjunction takes place at the common node or 90° from the common node of the two orbits. Owing to the fact that the two inclinations are sensibly the same, conjunctions which occur when the nodes of the two orbits on the fundamental plane are together will satisfy the mirror condition exactly, whatever the longitude of the conjunction point, since at this time the two orbital planes coincide. Now (**7**)

$$\theta_E = 328^\circ - 152^\circ{\cdot}7t$$

and

$$\theta_D = 276^\circ - 31^\circ{\cdot}0t,$$

where θ_E, θ_D are the longitudes of the nodes of Enceladus and Dione respectively, and t is measured in Julian years from the epoch 1889 April 0·5.

$\therefore$ the synodic period of the nodes = 2·96 years = B, say.

But the synodic period of the apse lines = 3·88 years = A, say.

$$\therefore 4B - 3A = 0^\circ{\cdot}20$$

or, within the accuracy of the derived elements (**14**),

$$4B = 3A.$$

A libration in the longitudes of Enceladus and Dione of period $3A = 11{\cdot}6$ years is to be expected. According to the *Connaissance des Temps*, there is a libration in the longitude of Enceladus of amplitude $20'{\cdot}0$ and period 12·2 years, while the longitude of Dione has a libration of the same period, and amplitude $\sim 2'$. According to the *Berliner Astronomisches Jahrbuch* (**6**), the amplitudes are $14'$ and $1'$ respectively, and the period is 11·1 years. It is seen that, within the accuracy of the elements, the libration period $= 3A$, as predicted.

As a further check, it was found that, according to the first set of elements (**7**), a conjunction of the nodes and a conjunction of the lines of apses occurred at $t = -5{\cdot}5$ years, thus giving an *exact* mirror configuration at conjunction at this epoch and at all epochs separated from it by multiples of 11·6 years. The agreement is less satisfactory for the second set of elements (**6**), but it should be noted that, although the rates of retrogression of the nodes may be reasonably well determined, owing to the small inclinations the position of either node at a particular epoch is poorly determined.

We may conclude that, within the limits of accuracy of the derived elements, the system of Dione and Enceladus conforms to the pattern of the arguments of Part II.

(*c*) *Mimas and Tethys.*—The following elements have been given (**7**):—

	Mean Motion	Inclin.	Eccen.	Mass/primary
Mimas	381°·994 5	1° 36′·5	0·019 0	$(16\,340\,000)^{-1}$
Tethys	190°·697 95	1° 4′·36	0·000 0	$(\ 921\,500)^{-1}$

In this case, therefore, we have $\sin i_1 > e$, and we would expect the mirror configurations to be related primarily to the position of the common node of the two orbits.

For two orbits whose inclinations to the fundamental plane of the system are i_1 and i_2 $(i_1 > i_2)$, and whose nodes on the fundamental plane of the system have longitudes θ_1 and θ_2 respectively, it can be shown that the longitude of the common node θ_j is given by

$$\theta_j = \theta_1 + \zeta,$$

where

$$\tan\zeta = \frac{\sin(\theta_1 - \theta_2)}{\tan i_1 \,.\, \cot i_2 - \cos(\theta_1 - \theta_2)}, \tag{8.01}$$

θ_j being measured along the fundamental plane of the system.

As a first approximation to the actual situation of Mimas and Tethys, we will consider that Tethys moves undisturbed by Mimas, that the two orbits are equally inclined to the fundamental plane, and that both orbits have zero eccentricity. Then, from (8.01),

$$\zeta = 90^\circ - \frac{(\theta_1 - \theta_2)}{2}.$$

If, therefore, conjunctions of Mimas and Tethys follow case (ii) of Section 4, the longitude of conjunction must always be $(\theta_M - \theta_T)/2$. If $\dot\theta_M, \dot\theta_T$ are the rates of revolution of the nodes of Mimas and Tethys on the fundamental plane, then, for this condition to hold, it is necessary that

$$2n_T - n_M = (\dot\theta_M + \dot\theta_T)/2$$

or

$$4n_T - 2n_M - \dot\theta_M - \dot\theta_T = E = 0. \tag{8.02}$$

But this is precisely the situation for these two bodies, as noted by H. Struve (**11**). Struve also noted that, if the mean longitudes (i.e. ignoring libration

terms) of the satellites are l_M and l_T, then

$$4l_T - 2l_M - \theta_T - \theta_M = 0. \tag{8.03}$$

Struve was thus led to enunciate the theorem that, ignoring libration, conjunctions of Mimas and Tethys occur always at a longitude which is the mean of the longitudes of the two bodies. Within the framework of the approximation considered here, this statement is equivalent to the statement that, ignoring libration, conjunctions occur always 90° from the common node of the two orbits. The mirror condition would, therefore, be satisfied at every conjunction.

In passing, it should be noted that the improvement of the determination of the elements of the two orbits with time has rendered (8.02) more exact, viz.

	Tisserand	*C. des T.* (7)	*B.A.J.* (6)
$E =$	$+3^\circ$	$+0^\circ{\cdot}532$	$+0^\circ{\cdot}254$

A second approximation to the actual problem is obtained by considering both orbits to be circular, but of different inclinations to the fundamental plane. Then the mean conjunction point is only *exactly* 90° from the common node if $\theta_M = 180^\circ + \theta_T$. Allowing for the different inclinations, we would expect a libration in mean longitude of both bodies of a period equal to an integral number of synodic periods of the nodes. A libration of amplitude $\sim 45^\circ$ and period ~ 70 years is present in the longitude of Mimas, and a small libration of the same period in the longitude of Tethys. For the satisfaction of the mirror condition at each phase of zero libration, on the present approximation, the period of libration, P, must be such that

$$(\dot\theta_M - \dot\theta_T)\, P/2 = 360^\circ m_1$$

or

$$P = 720^\circ m_1/(\dot\theta_M - \dot\theta_T), \tag{8.04}$$

where m_1 is an integer. At the same time, the phase of the libration must be such that zero libration corresponds to an opposition of the nodes.

With the elements of the *Connaissance des Temps* (7), an adjustment of 0·37 years in the libration period, and of 0·14 years in the epoch of zero libration would suffice to bring agreement with (8·04) with $m_1 = 29$. A comparison of the various sets of elements suggests that, owing to the long period of the libration, such adjustments are tolerable, but at the same time, the observational errors of both period and epoch being of the order of the synodic period of the nodes, little significance can be attached to such an agreement.

It should be noted that, if the epochs of zero libration correspond to oppositions of the two nodes, then the epochs of maximum and minimum libration correspond to conjunctions of the nodes. At such epochs, the mutual inclination of the orbits is at its smallest value of $(i_1 - i_2)$, and whatever the longitude of conjunction of the two bodies, a reasonable satisfaction of the mirror condition would occur.

A further approximation to the actual case of Mimas and Tethys is obtained by considering the eccentricity of Mimas' orbit. In this case, it is necessary that the epochs of zero libration correspond to conjunctions of the bodies when Mimas is at either its perisaturnium or its aposaturnium. With the mean motions used in Paper I, Table II, the rate of retrogression of the conjunction point of the two bodies is 218°·77 per year. Therefore, if $\dot\omega_M$ is the rate of advance of the apse line of the orbit of Mimas, it is necessary that

$$(\dot\omega + 218^\circ{\cdot}77)P = 180^\circ m_2, \tag{8.05}$$

where m_2 is an integer. At the same time it is necessary that the position of

perisaturnium be such that $\omega_M = \theta_M \pm 90°$ at the epochs of zero libration, where ω_M is the longitude of the perisaturnium of Mimas.

While it is possible to derive a set of values for θ_M, $\dot{\theta}_T$, P and $\dot{\omega}_M$ which satisfy (8.02), (8.04) and (8.05), and which are probably consistent with the observational accuracy of the determined values of these elements, it does not seem possible for ω_M to have the required value, unless there is an error $\sim 30°$ in the position of perisaturnium at the epoch of the published elements.

It may be concluded that the system of Mimas and Tethys agrees with the predictions of the arguments of Part II of the present paper to the first, and major, approximation, but it is uncertain whether the agreement is exact.

9. *Enforced eccentricities.*—As is well known, the occurrence of near-commensurability of mean motions of two bodies produces large inequalities, which may take the form of enforced eccentricities. The usual viewpoint is that the nature of the enforced eccentricities, in particular the rates of revolution of the apse lines, are consequences of the magnitudes of the mean motions. The viewpoint of the present paper is in no way in conflict with this traditional view, but it suggests that the mean motions have their observed values because, in such orbits, the enforced eccentricities will be such as to give the frequent satisfactions of the mirror condition. As is seen from the theorem of Newcomb (**15, 16**):

"In the case of two orbits, originally circular, if the movement of the point of conjunction, supposed unique, is small compared with the movements of the two satellites, the perturbations caused by one of the bodies engenders in the other an eccentricity, and the apse line of each orbit passes always through the point of conjunction."

The theorem thus states that, for near-commensurable mean motions, the enforced eccentricities will provide frequent satisfactions of the mirror condition.

10. *Jupiter's system.*—In this system, there are two near-commensurabilities where the ratio of integers is of the form $(A_2 + 1)/A_2$, namely J II/J I and J III/J II. From Table III of Paper I, we have

$$n_2/n_1 = \tfrac{1}{2} - 0{\cdot}001\,82$$

and

$$n_3/n_2 = \tfrac{1}{2} - 0{\cdot}003\,65.$$

In addition, there are the Laplace relations

$$n_1 - 3n_2 + 2n_3 = 0, \qquad (10.01)$$

$$l_1 - 3l_2 + 2l_3 = 180° \qquad (10.02)$$

where l denotes the mean longitude (**12**). The orbits have small inclinations and small but variable eccentricities.

The relations (10.01) and (10.02) mean that when J II and J III are in conjunction, J I is in opposition to them, giving a close approximation to the mirror condition which is repeated every 7·0509 days, which is less than a single period of J III.*

11. *Uranus' system.*—The satellite system of Uranus is unique among the satellite systems of the planets in that the orbits of the satellites are sensibly circular and coplanar, as is shown by the extensive investigation of Newcomb (**13**), who, in giving the elements, remarked that there is but slight evidence

* According to Griffin (**17**), whenever J II is in conjunction with J I or J III, the inner of the pair is near perijove and the outer of the pair near apojove, which would give an even closer approximation to mirror conditions.

of any real eccentricity of the orbits, and no evidence of any mutual inclinations. Thus, within the framework of the arguments of this paper, *for pairs of satellites* the system of Uranus presents a completely degenerate case, in that any conjunction of a pair of bodies therein will satisfy the mirror condition to a high degree of approximation. We would thus expect to find *no* preference for near-commensurability of mean motion among *pairs* of satellites of Uranus. Reference to Table IV of Paper I shows that this expectation is, indeed, confirmed. This may be considered evidence in support of the given interpretation of the preference for near-commensurability of mean motion found in the other satellite systems and the planetary system.

12. *Griffin's analysis.*—After the above investigation had been completed, the attention of the authors was drawn to an investigation of periodic orbits by Griffin (**17**). For the restricted case of n finite satellites moving in almost-circular, coplanar orbits, Griffin deduced his "symmetry theorem":

> "If a symmetrical conjunction occurs at any instant $t = t_0$, then the orbit of each satellite before and after the conjunction is symmetrical, both with regard to geometric equality of figures and with regard to intervals of time",

where a "symmetrical conjunction" is the coplanar case of a "mirror configuration". Griffin's "symmetry theorem" is thus a special case of the general "mirror theorem" deduced above.

Griffin further deduced that commensurabilities of the type

$$n_1/n_2 = (A_1 + 1)/A_1 \qquad (12.01)$$

are *excluded* for periodic orbits. This is in conflict, not only with the detailed conclusions of the present paper, but also with the actual occurrence of commensurabilities in the planetary and satellite systems, as is seen at once on reference to Tables I–IV of Paper I.

The contradiction is resolved by an examination of the basis of Griffin's analysis. He begins with a system of infinitesimal satellites moving in circular orbits, the periodicity of configurations being assured by the introduction of commensurabilities into these original orbits. He then considers the orbits that are obtained when the masses of the satellites are increased from zero, and obtains a series of periodic, nearly-circular orbits which "grow out of"* the original circular orbits. He shows that in the special case (12.01), periodic orbits do *not* grow out of the original circular ones.

His conclusion is thus that there is no continuity, with varying satellite masses, between the finite and the infinitesimal cases when (12.01) holds. His analysis does *not* preclude the situation whereby the perturbations of finite satellites obeying (12.01) build up into a pair of periodic elliptical orbits, beyond the limits of his approximation. The theorem of Newcomb cited in Section 9 is an example of this process. Griffin specifically points out that any commensurabilities of the form (12.01) can only yield periodic orbits if the orbits are elliptical, and the orbits would suffer large perturbations.

As evidence of the non-existence of periodic orbits obeying (12.01), Griffin cites the asteroid gaps. This interpretation of the asteroid gaps has been criticized, after Brown and Shook's analysis, in Section 7 of Paper I.

* This term is Griffin's.

13. *Conclusions.*—The mirror theorem in its two aspects, together with the periodicity theorem, provides reasons why near-commensurabilities of the form $(A_2+1)/A_2$ should be preferred among the mean motions of the planets and satellites, except in the case of coplanar, circular orbits, when they should be absent, in agreement with observation. It appears that such orbits satisfy Poincaré's conditions of stability if only that for a considerable part of the system's lifetime the bodies will remain near these orbits and suffer only librations.

The frequent occurrence of mirror configurations is most likely to occur in a system of nearly-circular, nearly-coplanar orbits, when it is satisfied at every conjunction and opposition (for a pair of bodies only). An extension of the arguments of this paper suggests, therefore, that the most stable configuration of orbits for a gravitating system of point-masses (i.e. neglecting tidal friction) is a system of nearly-circular, nearly-coplanar orbits, and that any gravitating system of bodies will spend most of its lifetime in such a configuration, whatever the original distribution of orbital inclinations and eccentricities may have been. The preference in the solar system for orbits of small inclinations and small eccentricities may be further evidence for the arguments of this paper.

The authors realize that these arguments are by no means rigorous (excepting the proofs of the mirror theorem and the periodicity theorem given in Sections 3, 4 and 5, which are rigorous) but they hope that the support given to these arguments by the observational evidence will encourage more interest in this problem of celestial mechanics. It is hoped that future work will show whether or not the other cases of near-commensurability of mean motions in the solar system fall within the pattern of the arguments of the present paper.

Acknowledgments.—It is a pleasure to acknowledge the encouragement of Professor W. M. Smart, and our debt to Professor G. R. Goldsbrough for drawing our attention to Griffin's analysis and for making valuable comments on the first draft of this paper. One of us (A. E. R.) is indebted to the Royal Society of Edinburgh, Robert Cormack Bequest Committee, for the award of a Research Fellowship.

University Observatory,
Glasgow, W.2:
1954 November 20.

References

(1) A. E. Roy and M. W. Ovenden, *M.N.*, **114**, 232, 1954.
(2) S. Herrick, *P.A.S.P.*, **64**, 237, 1952.
(3) G. H. Darwin, *Scientific Papers*, Vol. IV, Cambridge, 1911.
(4) W. M. Smart, *Celestial Mechanics*, London, 1953.
(5) S. Newcomb, *Astronomical Papers*, Vol. I, 8, Washington, 1882.
(6) *Berliner Astronomisches Jahrbuch für* 1938, p. 362.
(7) *Connaissance des Temps*, XXX, Paris, 1954.
(8) J. Woltjer, Jnr., *Ann. v. d. Sterr. te Leiden*, 1928.
(9) S. Newcomb, *Astronomical Papers*, Vol. III, 347, Washington, 1883.
(10) F. Tisserand, *Traité de Mécanique Céleste*, t. IV, 138, Paris, 1896.
(11) *Ibid.*, p. 129.
(12) *Ibid.*, p. 83.
(13) S. Newcomb, *Washington Astronomical Observations*, Appendix 1, 1873.
(14) H. Jeffreys, *M.N.*, **113**, 81, 1953.
(15) S. Newcomb, *Astronomical Journal*, VIII, No. 182.
(16) F. Tisserand, *loc. cit.*, p. 128.
(17) F. L. Griffin, *Periodic Orbits* (edit. F. R. Moulton), Carnegie Institution of Washington Publication, No. 161, p. 425, Washington, 1920.

THE INSTABILITY OF THE EINSTEIN UNIVERSE

W. B. Bonnor

(Communicated by A. G. Walker)

(Received 1954 December 9)

Summary

The stability of the Einstein universe is investigated using the new boundary conditions at surfaces of discontinuity, discovered by O'Brien and Synge. On certain assumptions it is shown that the Einstein model may represent a universe with inhomogeneities, but that, in general, these must be static otherwise the model is unstable.

In the case of an instantaneous disturbance in the Einstein universe, the processes which cause instability must involve discontinuities in the pressure or its time derivatives. If the matter present satisfies an equation of state, no physically plausible disturbance of the static condition is possible, and the model is stable. If there is no equation of state, small disturbances involving changes in the pressure or its derivatives can occur, and will, in general, lead to instability, though in certain special cases condensations may begin to form without initial instability.

---

1. *Introduction.*—Between the years 1930 and 1935 there appeared several papers on the stability of the Einstein model of the universe. The interest in this problem arose mainly because it was thought that the present expansion of the universe might have started from a perturbation of the Einstein equilibrium state. Although the writers all agreed that the Einstein model is unstable in some way or other, there was no agreement on the nature of the process in the model which would cause instability, or on the conditions in which an expansion, rather than a contraction, would begin to take place. These questions were not settled, and for various reasons cosmologists turned their attention to the point-source models of general relativity. However, these models are not without their difficulties—such, for example, as that of the formation of condensations in them—which make it premature to rule out the disturbed Einstein model. It therefore seems worth while to reconsider the stability of the Einstein universe in the light of certain recent developments in the theory of relativity.

A study of the previous work on the stability of the Einstein model clearly shows that much of the discordance in the results is due to the use of different boundary conditions at surfaces of discontinuity. The usual procedure was to suppose some disturbance to occur in the model at a definite time, say $t=0$, and to try to find, under certain assumed conditions of discontinuity of the metric and of the pressure and density, what happens for $t>0$. As twenty years ago the question of boundary conditions to be used for such problems in general relativity had received no systematic attention, it is not surprising that different sets of conditions were used, and different conclusions reached.

Quite recently this subject of boundary conditions at surfaces of discontinuity in general relativity has been studied by O'Brien and Synge (1), and by

Lichnerowicz (**2**). By the use of boundary conditions obtained by these authors it is possible to give a rather more definite answer than formerly to the question of the stability of the Einstein universe. Of course, in a subject like general relativity, where the fundamental equations are so far removed from physical intuition, sets of boundary conditions must depend on certain mathematical assumptions, so that in the absence of direct experimental confirmation they are arbitrary to a certain extent, and must be used tentatively. However, since essentially the same sets of conditions have been reached by O'Brien and Synge and by Lichnerowicz, starting from plausible but different assumptions, and since, as I have pointed out elsewhere (**3, 4**), the results of applying them to certain problems seem highly realistic, it is reasonable to use them in studying the stability of the Einstein universe.

In this paper I shall use these conditions in two problems connected with the stability of the Einstein model. The first of these, treated in Section 3, is the one stated above, of finding what instantaneous disturbances are possible, and what effects they would have if they occurred. The second problem, dealt with in Section 4, is that of the existence of models similar to Einstein's in being on a large scale homogeneous and static, but containing local inhomogeneities and motions. Before discussing these problems I give a short historical survey of the subject in Section 2. The conclusions which I reach, given in Section 5, are different from widely expressed views on the subject. They are, roughly, that the Einstein model is not necessarily unstable, nor *essentially* homogeneous, but that instability would result from certain disturbances in it which involve changes in the pressure.

2. *Historical survey.*—Using co-moving coordinates, the line-element of the homogeneous cosmological models of general relativity may be written

$$ds^2 = -e^{g(t)}\left[\left(1-\frac{r^2}{R^2}\right)^{-1}dr^2 + r^2(d\theta^2+\sin^2\theta\, d\phi^2)\right] + dt^2, \tag{2.1}$$

where R is a constant which may be real, imaginary or infinite, and $g(t)$ is an undetermined function of the time in terms of which the proper-density ρ and the proper-pressure p of the matter in the model are given by

$$8\pi\rho = 3R^{-2}e^{-g} + \tfrac{3}{4}\dot{g}^2 - \Lambda, \tag{2.2}$$

$$8\pi p = -R^{-2}e^{-g} - \ddot{g} - \tfrac{3}{4}\dot{g}^2 + \Lambda. \tag{2.3}$$

Here $\cdot$ denotes differentiation with respect to t, and Λ is the cosmological constant. The special case, corresponding to a static model, in which $g(t)=0$ and R is real, is called the Einstein universe, and for this the line-element may be put in either of the forms:

$$ds^2 = -\left(1-\frac{r^2}{R^2}\right)^{-1}dr^2 - r^2(d\theta^2+\sin^2\theta\, d\phi^2) + dt^2, \tag{2.4}$$

$$ds^2 = -\left(1+\frac{r^2}{4R^2}\right)^{-2}(dr^2+r^2d\theta^2+r^2\sin^2\theta d\phi^2) + dt^2. \tag{2.5}$$

The statement that the Einstein universe is unstable was first made in 1930 by Eddington (**5**), whose argument was as follows. If R is real then $Re^{\frac{1}{2}g}(=a)$ is the radius of the model (2.1), and (2.2) and (2.3) give

$$\frac{3}{a}\frac{d^2a}{dt^2} = \Lambda - 4\pi(\rho+3p). \tag{2.6}$$

For an Einstein universe with $p=0$ we have $4\pi\rho=\Lambda$; but if in such a model there is a slight disturbance causing a decrease in density, it is clear from (2.6) that d^2a/dt^2 is positive, so that the radius starts to increase. Moreover, since this expansion will cause a further decrease in density, it will continue, so that the model is unstable. A similar argument may be used, with the same result, if the pressure is not zero; and evidently increases in p or ρ will, on this reasoning, cause instability due to contraction. Eddington did not give any clear indication how the initial change in the density would occur, but considered that the formation of condensations might be a cause of the instability.

Shortly afterwards the instability of the Einstein universe was studied in some detail by Lemaître (**6, 7**), who tried to find, in particular, the nature of the process which would cause instability. He concluded that the process required to start an expansion from the Einstein equilibrium state was a reduction of the pressure in the model, which he called a "stagnation". The formation of condensations, except in so far as it is accompanied by a change in pressure, has, according to Lemaître, no effect on stability.

At about the same time McCrea and McVittie (**8, 9, 10**) attacked the problem by considering the differences between the volume of the completely homogeneous Einstein universe and of various static models similar to the Einstein universe but containing condensations represented as point singularities. They eventually concluded that, to the first order of approximation, the presence of condensations does not affect the total volume of the model, and that their method was not suitable for judging the effect of condensations on stability.

Subsequent studies of the instability were made by Dingle (**11**), Tolman (**12**) and Sen (**13, 14**), with conflicting results. The method of these authors was of the type referred to in Section 1, and it is at this stage that it becomes clear that the results obtained must depend to a large extent on the boundary conditions used. It should, however, be remarked that the importance of such conditions had evidently been realized in the earlier work of Lemaître.

3. *The conditions for instability.*—In this section we shall consider the stability of the Einstein universe to a disturbance which takes place at a definite instant $t=0$, and which preserves isotropy of pressure and also spherical symmetry about the origin of coordinates. By a disturbance is meant an event which corresponds to discontinuities in the functions which characterize the model; that is to say, in p, ρ and the metric tensor g_{ik} (but not in Λ). I suppose that the discontinuities must conform with the boundary conditions of O'Brien and Synge (**1**), which are that at the boundary $t(=x_4)=0$, the following quantities are continuous:

$$g_{ik},\ \partial g_{\mu\nu}/\partial x_4,\ T_k{}^4,\ g_{i\mu}T_k{}^\mu - g_{k\mu}T_i{}^\mu, \tag{3.1}$$

where $i, k = 1, 2, 3, 4$ and $\mu, \nu = 1, 2, 3$, and $T_k{}^i$ is the usual energy tensor of general relativity.

The most general expression for a non-static universe, using co-moving coordinates and supposing spherical symmetry about the origin, but not isotropy for all observers, is

$$ds^2 = -e^\lambda dr^2 - e^\omega(d\theta^2 + \sin^2\theta d\phi^2) + e^\nu dt^2, \tag{3.2}$$

where λ, ω and ν are functions of r and t. In accordance with the assumptions of the previous paragraph, we suppose that the line-element has form (2.4) for $t<0$, and form (3.2) for $t>0$, and that the conditions (3.1) apply at $t=0$. The

problem is then to find the restrictions imposed on λ, ω and ν by the boundary conditions and by the field equations

$$G_k^i - \tfrac{1}{2}\delta_k^i(G - 2\Lambda) = -8\pi T_k^i. \tag{3.3}$$

Since co-moving coordinates are being used, and the pressure is supposed isotropic, the only non-zero components of T_k^i are

$$T_1^1 = T_2^2 = T_3^3 = -p, \quad T_4^4 = \rho,$$

where p and ρ are the proper-pressure and proper-density.

In a previous paper (3) (hereafter referred to as I) I considered this problem for similar disturbances in a model which for $t<0$ is the non-static homogeneous universe (2.1), and I proved two theorems which, adapted to the special case of an Einstein universe, are as follows.

Theorem 1. If for $t>0$ the matter present satisfies an equation of state $p=f(\rho)$, where f is a single-valued function, then (3.2) reduces to (2.1).

Theorem 2. If the pressure is continuous at $t=0$ and if all its time derivatives at $t=0+$ are zero, then for $t>0$ (3.2) reduces to (2.4).

I shall refer to (2.1) as a Friedmann universe. Certain differentiability conditions are required in the proof of these theorems. Sufficient conditions are that p, ρ and g_{ik} possess time derivatives of all orders at $t=0+$, and that these derivatives be at least twice-differentiable functions of r. We shall assume throughout this section that these conditions are satisfied.

From Theorem 2 it follows that *disturbances of the type envisaged here can arise only if there is some change in the Einstein model which affects the pressure or its time derivatives: otherwise, the model is completely stable.* Theorem 1 shows that *if such a change takes place, and if there is an equation of state for $t>0$, the model will start expanding or contracting, and will, in fact, become a Friedmann universe.*

If there are discontinuities in the pressure and its derivatives, the stability may be studied, as was done in I for Friedmann models, by finding the initial rate of change of an element of proper volume dv of (3.2), given by $dv = e^{\frac{1}{2}\lambda+\omega}\sin\theta\, d\theta d\phi dr$. From I it follows that

$$\left[\frac{\partial(\log dv)}{\partial s}\right]_+ = 0,$$

$$\left[\frac{\partial^2(\log dv)}{\partial s^2}\right]_+ = -12\pi(p_+ - p_-), \tag{3.4}$$

$$\left[\frac{\partial^3(\log dv)}{\partial s^3}\right]_+ = -12\pi\left(\frac{\partial p}{\partial t}\right)_+, \tag{3.5}$$

where s is the proper time of an observer at rest in the coordinate system of (3.2), and $+$ and $-$ refer to $t=0+$ and $t=0-$. From (3.4) it follows that an instantaneous decrease of pressure at $t=0$ causes an expansion. This was the result found by Lemaître, who called the decrease of pressure a "stagnation". We see also from (3.5) that the expansion will be reinforced if $\partial p/\partial t$ is initially negative. If we suppose that the matter in the model is such that the pressure will be further diminished by an expansion, (3.5) indicates the instability of the model to reductions in the pressure. Evidently, a similar argument indicates that an increase of pressure will cause the Einstein model to contract.

In reaching his result Lemaître used arguments of questionable physical significance, and his methods were strongly criticized by McCrea and McVittie (**10**). It is therefore interesting that it is possible to arrive at his result by a method which uses as its only essential assumption (apart from the spherical symmetry and isotropy of pressure) the boundary conditions (3.1).

It was shown in I that p_+ is not a function of r, but $\dot{p}_+$ may be. This implies that if there is a discontinuity in the pressure, it must be universal, but that the pressure could begin to change in some isolated region. If there is an equation of state for $t>0$, the universal discontinuity in pressure is necessary for instability to develop, because of the following argument. If there is instability, the result, by Theorem 1, is the Friedmann model (2.1); now the boundary conditions require

$$g_+ = \dot{g}_+ = 0, \qquad \rho_+ = \rho_-,$$

so that from (2.3)

$$8\pi(p_+ - p_-) = -\ddot{g}_+.$$

If the pressure is continuous, $p_+ = p_-$ so that $\ddot{g}_+ = 0$. But owing to the equation of state, we then have

$$8\pi\left(\frac{\partial p}{\partial t}\right)_+ = 8\pi\frac{df}{d\rho}\left(\frac{\partial \rho}{\partial t}\right)_+ = \frac{df}{d\rho}[-3R^{-2}e^{-g_+}\dot{g}_+ + \tfrac{3}{2}\dot{g}_+\ddot{g}_+] = 0. \tag{3.6}$$

Further, differentiating (2.3) we obtain

$$8\pi\left(\frac{\partial p}{\partial t}\right)_+ = R^{-2}e^{-g_+}\dot{g}_+ - \dddot{g}_+ - \tfrac{3}{2}\dot{g}_+\ddot{g}_+. \tag{3.7}$$

Equations (3.6) and (3.7) give

$$\dddot{g}_+ = 0.$$

In this way it may be shown that $(d^n g/dt^n)_+ = 0$ for all n, so that (2.1) reduces to the Einstein model and there is no instability.

Unless, therefore, we are prepared to suppose that there is a universal discontinuity in the pressure, we have to rule out instability when there is an equation of state for $t>0$. The most realistic possibility for the development of instability seems to be the case where there is no such equation, and where the pressure is continuous, but $\partial p/\partial t$ (or a higher derivative) is discontinuous in some small region of space. In this case expansion or contraction starts according to (3.5). This case will be studied in the next section.

4. *Approximate Einstein models.*—In the previous section I have given a solution of the problem of the stability to certain types of disturbance of the Einstein model considered as completely homogeneous and static. One would naturally wish to interpret the homogeneity as the result of averaging the density and pressure over regions of a certain size, because inhomogeneities might be expected in a physical realization of the model. But for this interpretation to be permissible one must be sure that inhomogeneities could in fact exist in the Einstein universe. Stated rather more precisely, the question is whether there are solutions of the field equations representing static, inhomogeneous models which are similar to the Einstein model on a sufficiently large scale. If such solutions do not exist, then we are not entitled to regard the Einstein universe as a convenient model for something more complicated than it actually represents: in this case we must appreciate that any physical realization of the Einstein model must be strictly homogeneous, at least on a macroscopic scale.

Such static, inhomogeneous solutions do indeed exist; some have been published already, and I shall prove the existence of others later. We are therefore allowed to think of condensations in the Einstein universe without being in danger of any logical contradiction.

I have dealt with this point at some length because some writers on the Einstein universe seem to think that it does not allow the presence of condensations at all. There has been a tendency to confuse the existence of condensations with the formation of condensations. The latter is a different question, involving non-static solutions of the field equations.

The formation of condensations may be regarded from two points of view. First, we may consider processes in the strict Einstein model involving changes of the pressure and its derivatives, discussed in Section 3, which lead to changes in density and metric. Secondly, we may suppose that we *start* with a non-static model which resembles the Einstein model in the large, but which includes regions where density changes are taking place. In both cases we want to know whether the process of condensation can continue in bounded regions without causing the model to expand or contract as a whole. But in the second case we do not have to satisfy initial conditions at $t=0$, because the model is not supposed to start from the strict Einstein state.

To study these problems I consider what I call "approximate Einstein models". In these the variations from the Einstein metric are either confined to small regions of space, or (if not) are small except over limited regions.

Several approximate Einstein models were obtained by McCrea and McVittie (**8, 9, 10**), who found line-elements corresponding to Einstein universes containing particles represented as point-singularities. As an example, a special case of a solution of McVittie (**15**) may be quoted:

$$ds^2 = \left\{\frac{1-(m/2r)[1+(r^2/4R^2)]^{1/2}}{1+(m/2r)[1+(r^2/4R^2)]^{1/2}}\right\}^2 dt^2 - \frac{\{1+(m/2r)[1+(r^2/4R^2)]^{1/2}\}^4}{[1+(r^2/4R^2)]^2} \times (dr^2+r^2d\theta^2+r^2\sin^2\theta d\phi^2). \tag{4.1}$$

If $m=0$ this reduces to the Einstein model in the form (2.5), and if $R=\infty$ it gives the Schwarzschild metric in isotropic form. The pressure and density, obtainable from expressions given by McVittie, are functions of Λ, R, m and r, and we may not put $p=0$ without abolishing m.

Models such as (4.1) suggest that inhomogeneities could exist in an Einstein universe. This is the conclusion also from the work of Raychaudhuri (**16**), who found that there are exact static solutions of the field equations corresponding to an Einstein model containing spherical pockets of radiation of variable density. Indeed, as we shall see later, models can be constructed in which static, spherically symmetric regions of arbitrary density are embedded in an Einstein model.

In the approximate Einstein models which we shall consider, we suppose that the inhomogeneity, in which there may be motion, has spherical symmetry about the origin of coordinates, and lies within the sphere $r=a$ (constant). We assume that co-moving coordinates may be used so that the line-element has the form (3.2) inside $r=a$, and also that the pressure is isotropic. Outside $r=a$ the line-element is supposed to be that of the Einstein model (2.4), and across the boundary $r=a$, we suppose that the conditions (3.1) are satisfied with the obvious change that since the boundary is now $x_1=$ constant, the quantities which must be continuous are

$$g_{ik},\ \partial g_{\mu\nu}/\partial x_1,\ T_k^i,\ g_{i\mu}T_k^\mu - g_{k\mu}T_i^\mu, \tag{4.2}$$

where $i, k = 1, 2, 3, 4$ and $\mu, \nu = 2, 3, 4$. Evidently the model may contain any number of non-overlapping spheres of this type, each corresponding to an inhomogeneity. This model is somewhat similar to one used by Lemaître to study condensations in the expanding universe.

The conditions (4.2) give the following restrictions on $\lambda(r, t)$, $\omega(r, t)$ and $\nu(r, t)$ in (3.2):

$$\left.\begin{aligned} &\lambda(a, t) = -\log\left(1 - \frac{a^2}{R^2}\right), \quad \omega(a, t) = 2\log a, \quad \nu(a, t) = 0, \\ &\frac{\partial\omega}{\partial r}(a, t) = \frac{2}{a}, \qquad\qquad \frac{\partial\nu}{\partial r}(a, t) = 0. \end{aligned}\right\} \tag{4.3}$$

The continuity of p, required by (4.2), is secured by (4.3).

Let us consider first the case of a purely static inhomogeneity. Then we may take (3.2) as

$$ds^2 = -e^{\lambda}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + e^{\nu}dt^2,$$

where λ and ν are functions of r only. The field equations (3.3) give (Tolman (**17**))

$$8\pi T_1^{\ 1} = -8\pi p = -e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) + \frac{1}{r^2} - \Lambda, \tag{4.4}$$

$$8\pi T_2^{\ 2} = -8\pi p = -e^{-\lambda}\left(\frac{1}{2}\nu'' - \frac{1}{4}\nu'\lambda' + \frac{1}{4}\nu'^2 + \frac{\nu' - \lambda'}{2r}\right) - \Lambda, \tag{4.5}$$

$$8\pi T_4^{\ 4} = 8\pi\rho = e^{-\lambda}\left(\frac{\lambda'}{r} - \frac{1}{r^2}\right) + \frac{1}{r^2} - \Lambda, \tag{4.6}$$

where $'$ denotes d/dr. Equation (4.6) may be integrated to give

$$e^{-\lambda} = 1 - \frac{1}{3}\Lambda r^2 - \frac{8\pi}{r}g(r), \tag{4.7}$$

where

$$g(r) = \int^r \rho r^2 dr + K,$$

the density ρ being an arbitrary function of r, and the constant of integration K being chosen so that g satisfies the following, which is required by (4.3):

$$g(a) = \frac{a^3}{8\pi}\left(\frac{1}{R^2} - \frac{1}{3}\Lambda\right).$$

Equating the right-hand sides of (4.4) and (4.5), and substituting for λ from (4.7), we find a second-order differential equation for ν. It is clear that this admits a solution $\nu(r)$ containing two arbitrary constants which may be chosen to satisfy

$$\nu(a) = \nu'(a) = 0$$

without further restriction on $g(r)$. Hence for an inhomogeneity with an arbitrary density distribution there exists a solution satisfying the boundary conditions at $r = a$. *This leaves no doubt that static inhomogeneities can exist in the Einstein universe*, though they may not all be non-singular.

One further point about the static solutions is worth noticing. If p is constant for $r < a$, it follows from the identity

$$dp/dr = -\tfrac{1}{2}\nu'(p + \rho)$$

that $\nu' \equiv 0$ and the solution reduces to the homogeneous Einstein metric. In particular, in a model with $p = 0$, static inhomogeneities cannot exist.

Turning to non-static inhomogeneities, we face a much more difficult problem because it is necessary to integrate the field equations obtained from (3.2).

I shall not attempt the complete solution of this, but confine attention to the approximate solution which arises by supposing that the metric (3.2) inside the inhomogeneity differs only slightly from the Einstein metric (2.4).

First, let us suppose that ν, $\partial\lambda/\partial t$, $\partial\omega/\partial t$ and all their derivatives are small quantities of the first order. Then, neglecting small quantities of order higher than the first, the field equations for (3.2) (which are given in I) reduce to the following:

$$-8\pi T_1^{\ 1} = 8\pi p = -e^{-\omega} + e^{-\lambda}(\tfrac{1}{2}\omega'\nu' + \tfrac{1}{4}\omega'^2) - \ddot{\omega} + \Lambda, \tag{4.8}$$

$$-8\pi T_2^{\ 2} = 8\pi p = e^{-\lambda}(\tfrac{1}{2}\omega'' + \tfrac{1}{4}\omega'^2 + \tfrac{1}{2}\nu'' - \tfrac{1}{4}\omega'\lambda' - \tfrac{1}{4}\nu'\lambda' + \tfrac{1}{4}\nu'\omega') - \tfrac{1}{2}(\ddot{\lambda} + \ddot{\omega}) + \Lambda, \tag{4.9}$$

$$-8\pi T_4^{\ 4} = -8\pi\rho = -e^{-\omega} + e^{-\lambda}(\omega'' + \tfrac{3}{4}\omega'^2 - \tfrac{1}{2}\omega'\lambda') + \Lambda, \tag{4.10}$$

$$-8\pi T_4^{\ 1} = 0 = -\dot{\omega}' - \tfrac{1}{2}\omega'(\dot{\omega} - \dot{\lambda}), \tag{4.11}$$

where $\cdot$ denotes $\partial/\partial t$, and $'$ denotes $\partial/\partial r$.

Integrating (4.11), we have

$$e^{\lambda} = e^{\omega}\omega'^2 h(r), \tag{4.12}$$

where $h(r)$ is an arbitrary function, which for convenience we give the value

$$h(r) = \frac{1}{4}\left(1 - \frac{r^2}{R^2}\right)^{-1} e^{k(r)} \tag{4.13}$$

where $k(r)$ is another arbitrary function.

We now suppose further that

$$e^{\omega} = r^2[1 + \alpha(r, t)], \tag{4.14}$$

where α and its derivatives are small quantities of the first order. Then (4.12), (4.13) and (4.14) give, to first order,

$$e^{\lambda} = \left(1 - \frac{r^2}{R^2}\right)^{-1}(1 + \alpha + r\alpha' + k). \tag{4.15}$$

With (4.14) and (4.15), equations (4.8), (4.9) and (4.10) give after some calculation

$$-8\pi T_1^{\ 1} = 8\pi p = \Lambda - \frac{1}{R^2} - \ddot{\alpha} + \frac{\alpha}{R^2} + \frac{\nu'}{r}\left(1 - \frac{r^2}{R^2}\right) - \frac{k}{r^2}\left(1 - \frac{r^2}{R^2}\right), \tag{4.16}$$

$$-8\pi T_2^{\ 2} = 8\pi p = \Lambda - \frac{1}{R^2} - \ddot{\alpha} + \frac{\alpha}{R^2} + \frac{1}{2}\left(1 - \frac{r^2}{R^2}\right)\left(\nu'' + \frac{\nu'}{r}\right) + \frac{1}{2}\frac{r}{R^2}(\alpha' - \nu') + \frac{k}{R^2} - \frac{1}{2}\left(1 - \frac{r^2}{R^2}\right)\frac{k'}{r} - \frac{1}{2}r\ddot{\alpha}', \tag{4.17}$$

$$-8\pi T_4^{\ 4} = -8\pi\rho = \Lambda - \frac{3}{R^2} + \frac{3\alpha}{R^2} + \frac{r\alpha'}{R^2} - \frac{k}{r^2}\left(1 - \frac{3r^2}{R^2}\right) - \frac{k'}{r}\left(1 - \frac{r^2}{R^2}\right). \tag{4.18}$$

Equating the right-hand sides of (4.16) and (4.17), we have

$$\ddot{\alpha}' + \frac{\nu'}{r^2} - \frac{\nu''}{r}\left(1 - \frac{r^2}{R^2}\right) - \frac{\alpha'}{R^2} - \frac{2k}{r^3} + \frac{k'}{r^2}\left(1 - \frac{r^2}{R^2}\right) = 0,$$

which gives, on integration,

$$\ddot{\alpha} - \frac{\alpha}{R^2} - \frac{\nu'}{r}\left(1 - \frac{r^2}{R^2}\right) - \frac{\nu}{R^2} + \frac{k}{r^2}\left(1 - \frac{r^2}{R^2}\right) = p(t). \tag{4.19}$$

The boundary conditions (4.3) require that α, ν and ν' vanish at $r = a$ for all t, and as the continuity of λ requires $k(a) = 0$, we deduce that $p(t)$ in (4.19) must be identically zero. If we put

$$\beta(r, t) = \alpha - \frac{kR^2}{r^2}\left(1 - \frac{r^2}{R^2}\right), \tag{4.20}$$

(4.19) may then be written

$$\ddot{\beta} - \frac{\beta}{R^2} = \frac{\nu}{R^2} + \frac{\nu'}{r}\left(1 - \frac{r^2}{R^2}\right), \tag{4.21}$$

and we find from (4.16) and (4.18)

$$8\pi(p - p_0) = -\frac{\nu}{R^2}, \tag{4.22}$$

$$8\pi(\rho - \rho_0) = -\frac{3\beta}{R^2} - \frac{r\beta'}{R^2}, \tag{4.23}$$

where p_0, ρ_0 are the values of p and ρ in the homogeneous part of the model.

Equations (4.22) and (4.23) give the pressure and density inside the inhomogeneity in terms of two functions ν and β, which have to satisfy (4.21) and also the following conditions at $r = a$:

$$\nu(a, t) = \nu'(a, t) = 0, \tag{4.24}$$

$$\beta(a, t) = 0, \quad \beta'(a, t) = \frac{R^2}{a^2}\left(1 - \frac{a^2}{R^2}\right)k'(a) = \text{constant}. \tag{4.25}$$

In deriving (4.25) we remember that $k(a) = 0$, but $k'(a)$ is an arbitrary constant. The conditions (4.24) and (4.25) require

$$p(a, t) = p_0, \qquad p'(a, t) = 0, \tag{4.26}$$

$$\rho(a, t) - \rho_0 = \text{constant}.$$

To investigate the solution further, additional assumptions are necessary. Let us suppose that inside the inhomogeneity matter obeys a single-valued equation of state $p = f(\rho)$ (excluding, for the moment, $p = \text{const.}$ and $\rho = \text{const.}$), that p and ρ have derivatives with respect to r of all orders for $r < a$ and left derivatives with respect to r of all orders at $r = a$, and that p, ρ and these derivatives are at least twice-differentiable functions of t. Using (4.22) and (4.23), we find from (4.21)

$$\ddot{\rho} - \frac{\rho - \rho_0}{R^2} = \frac{3(p - p_0)}{R^2} + \frac{p'}{r}\left(2 - \frac{3r^2}{R^2}\right) + p''\left(1 - \frac{r^2}{R^2}\right);$$

and from the equation of state we may write this as a partial differential equation for p in the form

$$p'' = F(r, p, p', \dot{p}, \ddot{p}), \tag{4.27}$$

where F is a known function. If in this we put $r = a$ and use (4.26) we find that p'' is constant at $r = a$; and differentiating (4.27) with respect to r, it then follows easily that p''' also is constant at $r = a$. In this way it may be shown that all the derivatives of p with respect to r are constant at $r = a$, so that p is not a function of the time immediately within $r = a$. Using the differentiability condition on p, we can use (4.27) to extend this result to any point inside $r = a$. Thus p does not depend on t for $r < a$, and it is easy to see that ν, ρ and β do not depend on t either. It can also be shown that if $\rho = \text{const.}$ the solution is not a function of t. *Thus given an equation of state (other than $p = \text{constant}$) and appropriate differentiability conditions of p and ρ, the inhomogeneity must be static.* This result would be expected from Section 3, in which it was shown that in an Einstein model with an equation of state any disturbance must be universal.

If there is no equation of state, we may suppose that we are given $\nu(r, t)$ (i.e. the pressure) for $r < a$. In this case the right-hand side of (4.21) is known,

and the equation may be integrated to give

$$\beta = A(r)e^{t/R} + B(r)e^{-t/R} + H(r, t). \tag{4.28}$$

where $A(r)$ and $B(r)$ are arbitrary subject to conditions at $r = a$, and $H(r, t)$ is a particular integral, known if ν is given. (This form of β applies also in the special case where $p = \text{const.} = p_0$, which gives $H(r, t) = 0$.) The functions A and B would be determined if we were given appropriate initial conditions (e.g. $\rho(r, 0)$ and $\dot{\rho}(r, 0)$).

From (4.28) it is clear that β will increase or decrease exponentially with time except in special cases, such as, for example, that in which A and H vanish. This means that the metric (3.2) for $r < a$ will depart more and more from the Einstein state, and the assumptions of our method of approximation will eventually cease to apply. Before, however, we conclude that this means instability, we must examine a little more closely what is implied if β is an increasing or decreasing function of the time.

From the assumption that for $r > a$ the metric is that of the Einstein universe, it follows that the inhomogeneity is necessarily confined within $r = a$. The effect on stability of the motion inside $r = a$ is to be judged by the changes in the proper radius and proper volume of the inhomogeneity, which will affect the proper radius and proper volume of the whole model. If the proper radius and volume of the inhomogeneity tend to infinity with time, then the whole model is unstable because of expansion; if these quantities tend to zero, the whole universe contracts, though in this case we cannot be certain that this contraction will continue after the volume of the inhomogeneity becomes zero. If, however, the change in proper radius and proper volume of the inhomogeneity is zero, or remains small, then there is no reason why the motion inside $r = a$ should cause instability.

The proper radius l and the proper volume v of the inhomogeneity are, to the first order of approximation,

$$l = \int_0^a e^{\lambda/2}\, dr = \int_0^a \left(1 - \frac{r^2}{R^2}\right)^{-1/2} (1 + \tfrac{1}{2}\beta + \tfrac{1}{2}r\beta')\, dr, \tag{4.29}$$

$$v = \int_0^a \int_0^\pi \int_0^{2\pi} e^{\lambda/2 + \mu} \sin\theta\, d\phi\, d\theta\, dr = 4\pi \int_0^a r^2 \left(1 - \frac{r^2}{R^2}\right)^{-1/2} (1 + \tfrac{3}{2}\beta + \tfrac{1}{2}r\beta')\, dr. \tag{4.30}$$

In these expressions $k(r)$ does not appear because the terms which involve it may be integrated, and then vanish on the assumption that $k(a) = 0$ (already made), and that $k(r)$ tends to zero with r more rapidly than r. It may be stated here that the physical significance of $k(r)$ is connected with the spatial geometry of the model for $r < a$. When the other unknown functions, corresponding to the density and pressure, are determined, this geometry is not completely fixed, being unspecified to the extent of the arbitrary function $k(r)$. The arbitrariness disappears if appropriate initial conditions are given, as we shall see later in an example.

The contributions to the integrals (4.29) and (4.30) which could make l and v change rapidly with time are those involving, through β and β', $A(r)e^{t/R}$ and $H(r, t)$. In general, whatever the form of H, the dependence on $Ae^{t/R}$ will make l and v change rapidly and cause instability. However, in certain special cases A and H may be such that l and v remain nearly constant in time, even though at certain points in $r < a$, β may become large. For example, this will be so if

$H=0$ (which implies $p=p_0$) and A satisfies

$$\int_0^a \left(1-\frac{r^2}{R^2}\right)^{-1/2}(A+rA')\,dr=0, \qquad \int_0^a \left(1-\frac{r^2}{R^2}\right)^{-1/2} r^2(3A+rA')\,dr=0,$$

with $A \not\equiv 0$. In other words, the inhomogeneity may remain roughly the same size, though the density (which depends on β) changes considerably at some points in it. Thus condensation, or rarefaction, may begin without causing the whole model to contract or expand, though this will be so only for these special cases, and even in these we cannot be sure (since our method of approximation breaks down when β becomes large) that the integrals will not eventually begin to change rapidly, thus making the model unstable.

The solution, which is given essentially by (4.20) and (4.28), depends on a function $\nu(r, t)$, whose prescription amounts to the specification of the pressure $p(r, t)$; on the functions $A(r)$, $B(r)$ which may be considered as determined by the specification of ρ and $\partial\rho/\partial t$ as functions of r at $t=0$; and also on $k(r)$, determined by the metric at $t=0$. We have not so far supposed the model to have originated in any particular way: we may indeed regard it simply as a refinement of the Einstein model which represents, in so far as it is stable, a universe static and homogeneous on a large scale, but containing local irregularities and motions. On the other hand, we may ask whether a model of this type could have originated from the strict Einstein universe by a perturbation of the kind described at the end of Section 3. The result of this enquiry, to which we now turn, may be expected to restrict the arbitrariness of the functions A, B and k.

In addition to the conditions (4.3) at $r=a$, we must now impose the following at $t=0$ (which result from (3.1)):

$$\left.\begin{aligned} &\lambda(r,0)=-\log\left(1-\frac{r^2}{R^2}\right), && \dot\lambda(r,0)=0,\\ &\omega(r,0)=2\log r, && \dot\omega(r,0)=0,\\ &\nu(r,0)=0. \end{aligned}\right\} \tag{4.31}$$

Equations (4.3), (4.14), (4.15), (4.20) and (4.31) together give

$$\left.\begin{aligned} &\alpha(r,0)=\dot\alpha(r,0)=0,\\ &k(r)=0,\\ &\alpha(a,t)=\alpha'(a,t)=0,\\ &\nu(a,t)=\nu'(a,t)=0,\\ &\beta(r,t)=\alpha(r,t). \end{aligned}\right\} \tag{4.32}$$

Using (4.28) for β, and putting $t=0$, we see that from (4.32)

$$\left.\begin{aligned} &\beta(r,0)=A(r)+B(r)+H(r,0)=0,\\ &\dot\beta(r,0)=R^{-1}[A(r)-B(r)]+\dot H(r,0)=0, \end{aligned}\right\} \tag{4.33}$$

which determine A and B since $H(r, t)$ may be supposed known as $\nu(r, t)$ is given. In addition, (4.32) requires that

$$\left.\begin{aligned} &A(a)e^{a/R}+B(a)e^{-a/R}+H(a,t)=0,\\ &A'(a)e^{a/R}+B'(a)e^{-a/R}+H'(a,t)=0. \end{aligned}\right\} \tag{4.34}$$

It is easy to see that a function $\nu(r, t)$ may be chosen so that the problem has a solution. For example, if we take

$$\nu=r^2(r-a)^3q(r)e^{-\alpha t},$$

where α is a positive constant and $q(r)$ is a polynomial, we get a non-singular solution satisfying all the conditions (4.33) and (4.34). The solution corresponds to a disturbance in the Einstein model which takes place inside $r=a$, and in which the pressure begins to change from the equilibrium value at $t=0$. Although the pressure difference dies away eventually and may be arbitrarily small at all times, it causes instability in the general case because (4.33) requires $A(r)\neq 0$ so that the integrals (4.29) and (4.30) will eventually change rapidly, causing the whole model to contract or expand. However, in special cases this model can correspond to condensation or rarefaction inside $r=a$ without initial instability because the constants in $q(r)$ can evidently be chosen so that l and v do not change rapidly with time, even though $A(r)\neq 0$.

I shall now summarize the results of this section. We have seen that exact static solutions of the field equations exist corresponding to an Einstein universe with inhomogeneities. If there is an equation of state (other than $p=$ const.) no non-static approximate models with inhomogeneities are possible, but if there is no such equation, non-static models exist and are in general unstable, though there are special cases in which motion inside $r=a$ can occur without instability, including some in which condensations may begin to form. If a non-static approximate model results at $t=0$ from some change in the time derivatives of p inside a region $r=a$ of the strict Einstein model, its future is completely determined given $p(r, t)$ for $t>0$, $r<a$, and in general it will be unstable no matter how small is the pressure difference in the homogeneous and inhomogeneous parts of the model; but in special cases, condensations can begin to form without instability.

5. *Conclusion.*—From Section 4 it seems that we are justified in regarding the homogeneity of the strict Einstein model (2.4) as a smoothed out representation of a static universe with inhomogeneities (except in the case of a completely pressure-free model). Its static character must, however, be interpreted more strictly because it appears that, except in particular cases, systematic motions are accompanied by instability. Speaking roughly, we may say that the Einstein universe is not *essentially* homogeneous, but it *is* static*, except for the special motions mentioned in Section 4, and also, possibly, for rotational motions, which have been excluded from this work by the assumption of spherical symmetry, and which may not be accompanied by instability.

If we accept this view of the Einstein universe, it is clear that instability will result from any process which introduces motion, subject to the exceptions referred to above. Except in the case $p=0$, such a process will not follow from the mere presence of condensations of matter, because static solutions corresponding to them exist. From Section 3 it appears that the processes concerned are those which involve discontinuities in the pressure or its derivatives, since in the absence of these no instantaneous disturbance can occur. However, if there is an equation of state after the disturbance, the only permitted discontinuity of the type we have considered is a universal discontinuity in the pressure itself, and this is not likely to occur. (If it did, a Friedmann model would result.) Thus a model with an equation of state would probably be stable because no change whatever could take place in it.

* I think it would be wrong to conclude from this that molecular motions contributing to the pressure are themselves sufficient to cause instability. In all ordinary cosmological investigations, including this one, p and ρ are macroscopic quantities, and the theory is not intended to apply at the molecular level.

In a model without an equation of state, instability could arise through a gradual pressure change (i.e. a discontinuity in the time derivatives of p at $t=0$) inside a limited region; such a disturbance,e ven if small, will in general lead to condensation or rarefaction, and will cause instability through an expansion or contraction of the region in which it occurs. However, in special cases, the disturbance could die out or lead to condensation or rarefaction without immediate instability.

A universal decrease in the pressure in an Einstein model (if such could occur) would initiate an expansion, and an increase would start a contraction. For a gradual pressure change in a limited region, whether there is an expansion or contraction depends on the sign of dl/dt and dv/dt, which in turn depends on the pressure distribution, and the initial distributions of ρ and $\partial\rho/\partial t$.

If the Einstein universe contained evolving stars, changes of the pressure of radiation would presumably be sufficient to cause instability. If, however, one pictures it as filled with uniform gas at low temperatures (Bondi, **18**), or as a static mixture of gas and condensed, non-radiating matter, this work suggests that it would be stable.

As the boundary conditions (3.1) are still to some extent tentative, it is worth while emphasizing how convincingly they seem to work in the problems dealt with here. All the functions which they leave arbitrary have a reasonable physical interpretation, and the data required to determine the solutions are very much what one would expect on classical grounds. This has certainly not been true of sets of boundary conditions used in some of the previous work on this subject.

Department of Applied Mathematics,
The University,
Liverpool:
1954 *December* 8.

References

(1) O'Brien, S. and Synge, J. L., "Jump Conditions at Discontinuities in General Relativity", *Communications of the Dublin Institute for Advanced Studies*, Series A, No. 9, 1952.
(2) Lichnerowicz, A., "Étude Mathématique des Théories Relativistes de la Gravitation et de l'Électromagnétisme", Cours professé au Collège de France, 1952–53 (roneotyped).
(3) Bonnor, W. B., *Zeits. f. Astrophysik*, **38**, 10, 1954.
(4) Bonnor, W. B., *Proceedings of the International Congress of Mathematicians* 1954, p. 325, North-Holland Publishing Co.
(5) Eddington, A. S., *M.N.*, **90**, 668, 1930.
(6) Lemaître, G., *M.N.*, **91**, 490, 1931.
(7) Lemaître, G., *Annales de la Société Scientifique de Bruxelles*, Série A, **53**, 51, 1933.
(8) McCrea, W. H. and McVittie, G. C., *M.N.*, **91**, 128, 1930.
(9) McVittie, G. C., *M.N.*, **91**, 274, 1931.
(10) McCrea, W. H. and McVittie, G. C., *M.N.*, **92**, 7, 1931.
(11) Dingle, H., *M.N.*, **94**, 134, 1933.
(12) Tolman, R. C., *Proc. Nat. Acad. Sci.*, **20**, 169, 1934.
(13) Sen, N. R., *Zeits. f. Astrophysik*, **9**, 215, 1935.
(14) Sen, N. R., *Zeits. f. Astrophysik*, **10**, 291, 1935.
(15) McVittie, G. C., *M.N.*, **93**, 325, 1933.
(16) Raychaudhuri, A. K., *Bulletin of the Calcutta Mathematical Society*, **44**, 31, 1952.
(17) Tolman, R. C., *Relativity, Thermodynamics and Cosmology*, p. 242, Clarendon Press, 1934.
(18) Bondi, H., *Cosmology*, p. 118, Cambridge, 1952.

THE INTRINSIC COLOURS OF GALACTIC CEPHEID VARIABLES

D. W. N. Stibbs

(Communicated by the Director, University Observatory, Oxford)

(Received 1955 February 24)

Summary

The intrinsic colours of type I galactic cepheids are determined from the photoelectric colours of Eggen, using a method in which the observed colours are corrected for reddening by means of a model of the reddening medium which gives the colour excesses of B stars as a function of distance and the galactic coordinates. Estimates of the distances and reddening of the cepheids are found from the model by converting colour excess into photographic absorption and solving the distance modulus equation using Baade's zero-point of the period–luminosity curve. Solutions by least squares are then made for the coefficients in linear relations for the intrinsic colours at maximum and minimum light for cepheids belonging to Eggen's types A, B and C. The following solution for the intrinsic colours on the **P, V** System is adopted :

Types A, B and C : $(\mathbf{P}\text{–}\mathbf{V})_{\max} = +0^{m}\cdot17 + 0^{m}\cdot18 \log P$

Types A and B : $(\mathbf{P}\text{–}\mathbf{V})_{\min} = +0^{m}\cdot08 + 0^{m}\cdot78 \log P$

Type C : $(\mathbf{P}\text{–}\mathbf{V})_{\min} = -0^{m}\cdot09 + 0^{m}\cdot77 \log P.$

It is shown that these intrinsic colours give colour excesses and distance moduli which support the hypothesis that cepheid distances and periods are not correlated. The intrinsic colours are much bluer than hitherto supposed, and at maximum light there is some similarity with the colours of cepheid variables in the Small Magellanic Cloud.

The intrinsic colours are applied to an examination of the connection between differential galactic rotation and the zero-point of the period–luminosity curve. It is found that Baade's zero-point gives an Oort constant $A = 20$ km/sec/kpc for a group of nearby cepheids for which the average absorption is about 2 magnitudes per kpc. The recovery of Joy's original value of A for the cepheids is shown to be due to the fact that the greater absorption required by the intrinsic colours compensates for the increase in Joy's distance moduli made necessary by Baade's revision of Shapley's zero-point of the period–luminosity curve.

Introduction.—The determination of the interstellar absorption of light by means of the observed colours of stars requires a knowledge of their unreddened (or intrinsic) colours, and a value for the ratio of the total photographic absorption to the selective absorption or colour excess. In the case of the B stars, there is an adequate sample of nearby objects in high galactic latitudes from which the first requirement, the intrinsic colours as a function of spectral type, can be obtained by supposing that the amount of reddening which their light has undergone is negligibly small (**1, 2**), or can be estimated (**3, 4, 5**). On the other hand, cepheid variables belonging to Baade's population type I are strongly concentrated in the galactic plane. They are, moreover, intrinsically more luminous than the B stars and thus can be seen at greater distances. Consequently it is unlikely that there are many cepheids whose observed colours can be considered as unreddened.

Attempts have been made to avoid the difficulty presented by the spatial distribution of the type I cepheids by an appeal to the colours of non-variable supergiant stars. The method involves two steps (i) the assignment of spectral types to the cepheids, and (ii) the determination of the intrinsic colours of supergiants as a function of spectral type. Originally (i) was carried out by means of a relation between period and median-spectrum, but this has been modified by more precise spectral classifications on the Yerkes System which lead to well-defined period-spectrum relations at maximum and minimum light (**6**). This is probably the most certain part of the method. In step (ii), in addition to having to decide upon how much reddening should be allowed for in assessing the intrinsic colours of supergiant stars in a variety of galactic latitudes, it is felt that some consideration should be given to the basic assumption that variable and non-variable supergiants of the same spectral type have the same intrinsic colour.

The method which refers to the colours of non-variable supergiants is discarded in this paper in favour of a more direct procedure, similar to that used in (**4**) for the determination of the intrinsic colours of the B stars, in which individual corrections for reddening are applied to the observed colours, the amount depending upon distance and galactic latitude. Earlier attempts (**7, 8**) to determine the intrinsic colours of cepheids in this way were based upon photographic observations and upon a cepheid distance-scale which has been considerably modified by Baade's revision of the zero-point of the period-luminosity curve. In the present work a determination of the intrinsic colours is made from photoelectric colours using the new cepheid distance-scale. Eggen's observations (**9**) of cepheid colours reduced to the **P, V** System (**10**) are used, thereby making it possible to compare the colours of galactic cepheid variables with the **P, V** colours of cepheids in the Small Magellanic Cloud (**11**).

1. *The method*

Let **P** and **V** be the photographic and photovisual magnitudes on the **P, V** System, **P**–**V** being the colour, and let us suppose that the cepheid colours at maximum and minimum light can be written in the form

$$\left.\begin{aligned}(\mathbf{P}\text{–}\mathbf{V})_{\max} &= a + b\log P + E,\\ (\mathbf{P}\text{–}\mathbf{V})_{\min} &= c + d\log P + E,\end{aligned}\right\} \qquad (1)$$

where P is the period and E, the colour excess, is some complicated function of distance and the galactic coordinates. The first two terms in these expressions are the intrinsic colours at maximum and minimum light. The intrinsic colour is assumed to be linear in log P since the colour amplitudes

$$\Delta(\mathbf{P}\text{–}\mathbf{V}) = (\mathbf{P}\text{–}\mathbf{V})_{\min} - (\mathbf{P}\text{–}\mathbf{V})_{\max}, \qquad (2)$$

which are probably unaffected by space reddening, were found by Eggen (**9**) to exhibit a strong linear correlation with log P.

Since the total photographic absorption A_{pg} due to selective absorption is proportional to the colour excess, i.e. $A_{pg} = \chi E$, we have, considering only one of the equations in (1),

$$\chi\{(\mathbf{P}\text{–}\mathbf{V}) - (a + b\log P)\} = A_{pg}. \qquad (3)$$

If A_{pg} for a given cepheid can be inferred from the colour excesses of nearby B stars on the system of Stebbins, Huffer and Whitford, we have $A_{pg} = \chi_1 E_1$ and equation (3) becomes

$$(\chi/\chi_1)\{(\mathbf{P}\text{–}\mathbf{V}) - (a + b\log P)\} = E_1. \qquad (4)$$

Now the photoelectric colours are probably sufficiently accurate to enable E_1 to be treated as the observational quantity in the equation of condition (4). In this case, it is legitimate to apply the ordinary method of least squares to solve for a, b and χ/χ_1. But the determination of χ/χ_1 is properly a photometric problem, and it would be more appropriate to adopt a value for the ratio χ/χ_1 thereby throwing all the weight of the solution into the determination of a and b.

In (**8**), equation (3) was used in the form

$$a+b\log P+cA_{pg}=(\mathbf{P}\text{–}\mathbf{V}),$$

and a least squares solution was made for a, b and c, using photographic observations of cepheid colours. This procedure is, however, only correct if A_{pg} is known accurately or with an error considerably smaller than the error in the colours, otherwise a multi-error least-squares solution would be more appropriate, if manageable. In the present work, the observed colours are corrected for reddening before applying the least squares method, the correction being derived from the colours of B stars.

If the colour excesses of the B stars can be approximated by a function involving the distance r, the galactic latitude b, and if possible the galactic longitude l, the formula for colour excess may be used to estimate the reddening experienced by a given cepheid. The correction for reddening is then $(\chi_1/\chi)E_1(r, l, b)$, or $(\chi_1/\chi)E_1(r, b)$ if the average over the observable sector of longitudes has to be taken. In order to establish a suitable interpolation formula for E_1, it is necessary to adopt a model which is capable of representing the averaged structure of the reddening medium.

2. *The reddening medium*

It is well known that the variation of stellar density with distance z from the galactic plane can be adequately represented by the relation $D(z)=D_0e^{-z/\beta}$. For the A stars alone, β is approximately 100 pc, but considerably greater values of β are found for later spectral types. Taking all the stars together, Oort's data (**12**) can be well represented for $z\leqslant 500$ parsecs by the simple exponential formula with $\beta=230$ pc. But from the appearance of extra-galactic nebulae seen edge-on, the obscuring matter might be expected to be more concentrated towards the galactic plane than the stars in general. The exponential distribution for the obscuring matter with $\beta=105$ pc was adopted by Williams (**13**).

The averaged density distribution of the reddening material could be determined from the colour excesses of the B stars if a curve showing the dependence of $E\sin b$ upon $z(=r\sin b)$ could be established from the data. Thus if $c(z)$ is the coefficient of selective absorption at height z above or below the galactic plane, we have

$$c(z)=dE\sin b/dz, \tag{5}$$

and the density of the reddening material, which is proportional to $c(z)$, follows by differentiation of the curve. With $c(z)=c_0e^{-z/\beta}$, the colour excess for an object at distance r in galactic latitude b is given by

$$\begin{aligned}E(r, b)&=c_0\int_0^r\exp\left(-\frac{r\sin b}{\beta}\right)dr\\&=\frac{c_0\beta}{\sin b}\left(1-\exp\left(-\frac{r\sin b}{\beta}\right)\right).\end{aligned} \tag{6}$$

This form for the reddening has been used extensively (**14, 15, 16, 17**).

In the present application, use is made of three independent determinations of the parameters c_0 and β, two of which give the colour excess distribution averaged over the sector of galactic longitude covered by the data on B star colours, the third giving the distribution within small zones bounded by circles of galactic latitude and longitude in each of which average values of the parameters are used.

2.1. *Reddening according to Oort.*—In Oort's work (**3**), a correction of $+0^{m}\cdot049$ was applied to the zero-point of the colour excesses of the B stars on the system of Stebbins, Huffer and Whitford (SHW). Using distances r calculated from secular parallaxes and the rotational effect in the radial velocities with $A = 19$ km/sec/kpc, Oort obtained a curve giving $E_1 \sin b$ as a function of z. This curve was differentiated by Parenago (**15**) who obtained

$$\log (dE_1 \sin b/dz) = -3\cdot42 - 0\cdot00356\, z.$$

This corresponds to the exponential distribution of $c(z)$ in equation (5) with $c_0 = 0^{m}\cdot38$ per kpc, and $\beta = 122$ pc. These values of the parameters give the general distribution of the reddening medium averaged over the longitude sector from $l = 0^{\circ}$ to 200°. The value of c_0 determined in this way is rather sensitive to the correction applied to the zero-point of the colour excesses.

2.2. *Reddening according to Parenago.*—Equation (6) written in the form

$$A_{pg} = \frac{a_0 \beta}{\sin b}\left(1 - \exp\left(-\frac{r \sin b}{\beta}\right)\right), \tag{7}$$

has been developed and applied extensively by Parenago in two papers (**15, 16**). In the first paper, the following mean values of the parameters were adopted for the whole sky: $a_0 = 3^{m}\cdot4$ per kpc, $\beta = 100$ pc. The second paper takes account of irregularities in the absorption, both in latitude and longitude. The determination of the parameters a_0 and β has been described by Parenago (**16**, p. 131) as follows:

> We determine for each star whose photoelectric colour index is known, for every place with counts of extra-galactic nebulae, etc., the values of A_{pg} and r. Assuming some initial value of β (that taken was $\beta = 100$ pc), we calculate for each determination of A_{pg} and r the value of a_0 from formula (7) and place it on a diagram as a function of l and b. Further, having taken account of the unavoidable errors in the determination of A_{pg} and r, we group our determinations of a_0 in such a way that without exceeding the limits of these errors, we determine regions on the diagram such that a_0 may be regarded as constant in each. In this way we get a whole series of such regions for each of which we can determine by the method of least squares the most probable a_0 and the correction to the assumed β.

Parenago found that within the limits of error of the determinations, the value of β was independent of the longitude, the mean value being $\beta = 100 \pm 4$ pc. Values of a_0 were assigned to 122 regions of which 81 lie within $\pm 10^{\circ}$ of the galactic equator. For regions close to the equator it is reasonable to suppose that most, if not all of the weight in the determination of the values of a_0 comes from the colours of B stars. Parenago used Florya's intrinsic colours (**4**, p. 12) which correct the zero-point of the SHW colour excesses by $+0^{m}\cdot04$, and obtained the absorption from $A_{pg} = 10E_1$. The distances of the B stars were calculated for adopted absolute magnitudes.

2.3. *Reddening according to van Rhijn.*—In an investigation on the differential absorption averaged over the longitude interval from $l = 320°$ to $210°$, van Rhijn (**17**) obtained

$$E_1(r, b) = \frac{0^m{\cdot}031}{\sin b}\,(1 - e^{-0{\cdot}0081 r \sin b}).$$

This corresponds to $c_0 = 0^m{\cdot}25$ per kpc and $\beta = 124$ pc in equation (6). Van Rhijn used the SHW colour excesses with a zero-point correction of $+0^m{\cdot}03$, as calculated in (**5**). The distances were derived from interstellar line intensities, the rotational effect in the radial velocities with $A = 18$ km/sec/kpc, secular parallaxes, and from the density law and luminosity function. It was shown that the difference between the observed mean colour excess E_1 and that found from the formula with distances estimated in the above mentioned ways was quite generally small. The formula was also shown to hold for the longitude sectors $320°$ to $110°$, and $110°$ to $210°$ separately, although derived from stars at all longitudes combined.

3. *Solutions for the intrinsic colours*

According to Eggen (**9**) cepheids may be divided into types A, B or C on the basis of relations between period, light-amplitude and colour-amplitude. The light-amplitudes of types A and B are strongly correlated with their periods, whereas the type C amplitudes are nearly uncorrelated. The colour-amplitudes, on the other hand, are strongly correlated with the periods for all three types, but the A and B types average about $0^m{\cdot}2$ greater in colour-amplitude than the C types. Following Eggen, we shall suppose that cepheids of types A, B and C at maximum light satisfy the same intrinsic colour relation. We then have to find two relations for the intrinsic colours at minimum light, one for types A and B, and another for type C. The solutions given in this Section are made from 37 cepheids of which 28 belong to types A and B, and 9 to type C. The intrinsic colours of type C cepheids at minimum light cannot be determined from the limited amount of data. The relation may, however, be inferred by adding the type C colour-amplitudes to the intrinsic colours of types A, B and C at maximum light.

The equation of condition for the intrinsic colours of the cepheids is

$$a + b \log P = (\mathbf{P}\text{–}\mathbf{V})_{corr},$$

where $(\mathbf{P}\text{–}\mathbf{V})_{corr}$, the observed colour corrected for space reddening, is given by

$$(\mathbf{P}\text{–}\mathbf{V})_{corr} = (\mathbf{P}\text{–}\mathbf{V})_{obs} - (\chi_1/\chi)E_1(r, b).$$

The reddening has been estimated by means of the distributions of colour excess in Section 2, $E_1(r, b)$ being calculated for each cepheid by successive approximations in the equation for the distance modulus

$$m - M = 5 \log r + \chi_1 E_1(r, b) - 5,$$

where M is the absolute magnitude from Shapley's period–luminosity curve (**18**) with a zero-point correction of $-1{\cdot}4$ magnitudes. Values for χ_1 and χ_1/χ are also required. We have adopted $\chi_1 = 9$ as confirmed by Whitford (**19**), and $\chi_1/\chi = 1{\cdot}9$ from the work of Seares and Joyner (**20**).

The solutions for the intrinsic colours are given in Table I. The root mean square errors of the coefficients in the linear relations are also given in the Table. The root mean square error of an equation of condition is $\pm 0^m{\cdot}14$ at maximum and minimum light for the solutions designated Oort and van Rhijn, in which the

TABLE I

Intrinsic colours of Cepheids at maximum and minimum light

Reddening adopted	Maximum	Minimum	
	Types A, B and C	Types A and B	Type C
Oort	$+0^m\cdot16+0^m\cdot08\log P$ $\pm\cdot07\quad\pm\cdot08$	$+0^m\cdot07+0^m\cdot68\log P$ $\pm\cdot09\quad\pm\cdot10$	$-0^m\cdot10+0^m\cdot67\log P$ $\pm\cdot07\quad\pm\cdot08$
Parenago	$+0^m\cdot17+0^m\cdot18\log P$ $\pm\cdot05\quad\pm\cdot05$	$+0^m\cdot08+0^m\cdot78\log P$ $\pm\cdot06\quad\pm\cdot07$	$-0^m\cdot09+0^m\cdot77\log P$ $\pm\cdot05\quad\pm\cdot05$
van Rhijn	$+0^m\cdot22+0^m\cdot12\log P$ $\pm\cdot07\quad\pm\cdot08$	$+0^m\cdot12+0^m\cdot74\log P$ $\pm\cdot08\quad\pm\cdot10$	$-0^m\cdot04+0^m\cdot71\log P$ $\pm\cdot07\quad\pm\cdot08$

mean colour excess distributions were used. This error is reduced by 36 per cent to $\pm0^m\cdot09$ in the Parenago solution which takes account of irregularities in the longitude distribution. Since the difference between the intrinsic colours in the solutions Parenago *minus* Oort, and Parenago *minus* van Rhijn are not statistically significant, we shall adopt the intrinsic colours according to the Parenago solution on account of its greater weight. Thus we adopt the following relations for the intrinsic colours of type I galactic cepheids

$$\left.\begin{aligned}&\text{Types A, B and C:}\quad (\mathbf{P}\text{–}\mathbf{V})_{max}=+0^m\cdot17+0^m\cdot18\log P,\\&\text{Types A and B:}\quad (\mathbf{P}\text{–}\mathbf{V})_{min}=+0^m\cdot08+0^m\cdot78\log P,\\&\text{Type C:}\quad (\mathbf{P}\text{–}\mathbf{V})_{min}=-0^m\cdot09+0^m\cdot77\log P.\end{aligned}\right\}\qquad(8)$$

The colours which lead to this solution are shown in Fig. 1.

4. *Some properties of the adopted solution*

Eggen (**9**) has given a set of relations for the intrinsic colours which were derived from eight cepheids. These are the bluest, for their periods, of the type I cepheids for which photoelectric observations of colour are available. Eggen's intrinsic colours on the **P, V** System are as follows (**21**):

$$\left.\begin{aligned}&\text{Types A, B and C:}\quad (\mathbf{P}\text{–}\mathbf{V})_{max}=-0^m\cdot04+0^m\cdot60\log P,\\&\text{Types A and B:}\quad (\mathbf{P}\text{–}\mathbf{V})_{min}=-0^m\cdot13+1^m\cdot22\log P,\\&\text{Type C:}\quad (\mathbf{P}\text{–}\mathbf{V})_{min}=-0^m\cdot32+1^m\cdot20\log P.\end{aligned}\right\}\qquad(9)$$

The relations (8) and (9) are compared in Fig. 2. The most striking feature of the comparison is the marked difference between the slopes of the lines in the two sets of intrinsic colours. The difference amounts to about seven or eight times the root mean square error in the present determination of the coefficients of $\log P$. The significance of this difference is now examined.

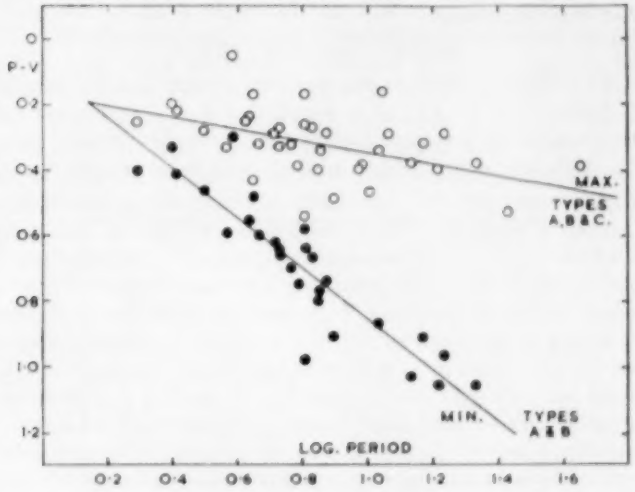


FIG. 1.—*The intrinsic colours of cepheid variables at maximum and minimum light according to the relations* (8). *The observed colours corrected for space reddening are shown by open and filled circles.*

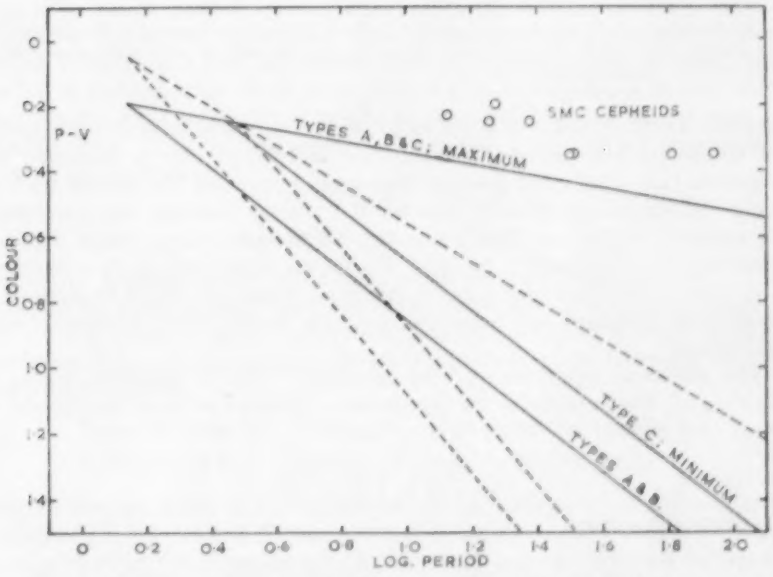


FIG. 2.—*Comparison between the adopted intrinsic colours (full lines) and Eggen's intrinsic colours (broken lines). Open circles give the observed colours at maximum light for cepheid variables in the Small Magellanic Cloud.*

4.1. *Correlation between distance and period.*—If the distances of the cepheids are calculated from the relation

$$\log r = 1 + \tfrac{1}{5}(m - M - \chi E), \tag{10}$$

where E is the difference between the observed colour and the intrinsic colour for a given period, and $\chi = 4{\cdot}75$, it is found that Eggen's intrinsic colours (9) give distances which show a strong correlation with the periods, as shown in Fig. 3. We wish to enquire if such a correlation is physically or statistically

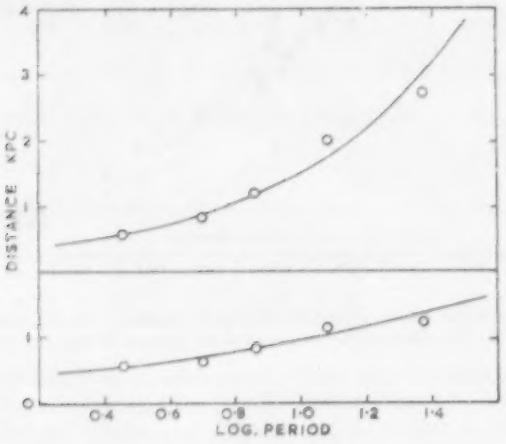


FIG. 3.—*Correlation between cepheid distances and periods. The linear regression of log r_{kpc} on log P is plotted in the figure as r_{kpc} on log P. The upper and lower panels give the correlation for distances based upon Eggen's intrinsic colours* (9), *and the adopted solution* (8) *respectively. Open circles denote group means.*

possible. On the physical side, the only known correlations involving the periods is a correlation between the z-distribution and $\log P$, due to Parenago and Grigorieva (**22**) who found that the concentration towards the galactic plane is greater for the longer periods than for the shorter periods, and the related conclusion of Shapley and McKibben (**23, 24**) that the longer-period cepheids favour regions of high stellar density. On the statistical side, we have to decide if the large correlation which Eggen's intrinsic colours give could be ascribed to chance if the distances and periods of cepheids in the parent population were uncorrelated.

The statistical significance of the correlation may be tested by means of Fisher's $z - \zeta$ distribution for the correlation coefficient which is approximately normal with mean

$$\frac{\rho}{2(n-1)}\left\{1 + \frac{1+\rho^2}{8(n-1)}\right\}$$

and standard deviation

$$\sigma \simeq \left\{\frac{1}{n-1} + \frac{4-\rho}{2(n-1)^2}\right\}^{1/2},$$

n being the number in the sample and ρ the correlation coefficient in the parent population. Consider first the case $\rho = 0$, that is when the distances and periods

of cepheids in general are uncorrelated. The correlation coefficient for the linear correlation of Eggen's values of $\log r$ with $\log P$ is $+0{\cdot}64$ from a sample of 38 stars.* The deviation in the $z-\zeta$ distribution is $4{\cdot}47\sigma$, and the probability of finding $\log r$ and $\log P$ positively correlated with a correlation coefficient $0{\cdot}64$ or greater is less than 1 in 2×10^5. Now the distances calculated by means of the intrinsic colours (8) give a correlation coefficient $+0{\cdot}36$, corresponding to a deviation in $z-\zeta$ of $2{\cdot}24\sigma$. In this case the probability is less than 1 in 80. Consider now the possibility that the cepheids have an artificial correlation between $\log r$ and $\log P$ due to observational selection. We can take account of this by supposing that the parent population itself is similarly correlated with a correlation coefficient ρ. If we suppose that ρ is as large as $0{\cdot}25$, we get a deviation in $z-\zeta$ of $2{\cdot}96\sigma$ in Eggen's case, and the probability is less than 1 in 650 that by chance alone we should find $\log r$ and $\log P$ positively correlated in the sample with a correlation coefficient $0{\cdot}64$. On the other hand, for distances computed with the help of equations (8) the deviation in $z-\zeta$ is $0{\cdot}69\sigma$, and the correlation found in the sample is entirely due to chance. In Eggen's case, however, the correlation is still not due to sampling, and the conclusion must be drawn that the intrinsic colours (9) which lead to such a large correlation are questionable.

Evidently the eight cepheids which Eggen originally adopted as unreddened are, in fact, appreciably reddened with increasing values of their periods. The reason for this can be traced to the coincidence that if the eight cepheids are arranged in order of increasing $\log P$, the last six are in order of increasing distance, ranging from the nearest at $0{\cdot}3$ kpc to the farthest at $4{\cdot}0$ kpc, and the last five are in order of decreasing galactic latitude. The spatial distribution is clearly favourable towards appreciable reddening for the longer periods.

4.2. *The colour-spectral type relation.*—The spectral classification of the cepheids on the Yerkes System has been shown by Code (**6**) to give period-spectrum relations in which the spectral types at maximum light are nearly the same for all periods, most, if not all of the well known progressive lateness with period shown by the spectral type at median phase being due to the variation at minimum light. Code's assignments of spectral type can be represented accurately by the relations

$$\left.\begin{aligned}\text{Sp. (max)} &= 4{\cdot}8+1{\cdot}5\log P,\\ \text{Sp. (min)} &= 3{\cdot}0+12{\cdot}3\log P,\end{aligned}\right\} \qquad (11)$$

in which the spectral type is in sub-class units measured from type Fo. The relation for Sp. (min) has only been calculated for cepheids belonging to Eggen's types A and B since the type C spectra probably obey a slightly different relation at minimum light. From equations (8) and (11) we obtain the following colour-spectral type relations:

$$\left.\begin{aligned}&\text{Types A, B and C:}\quad (\mathbf{P}\text{–}\mathbf{V})_{\max} = -0{\cdot}40+0{\cdot}12\ \text{Sp. (max)},\\ &\text{Types A and B:}\quad\ (\mathbf{P}\text{–}\mathbf{V})_{\min} = -0{\cdot}11+0{\cdot}063\,\text{Sp. (min)}.\end{aligned}\right\} \qquad (12)$$

These relations are shown in Fig. 4 with the intrinsic colours of selected cepheids for which **P, V** colours and Yerkes spectral classifications are available. The colours plotted in the diagram are the observed colours corrected for space reddening as in Fig. 1.

* One of this number, with $\log P=0{\cdot}83$, was omitted in the solutions in Section 2 as it gave a colour residual greater than three times the root mean square error of an equation of condition.

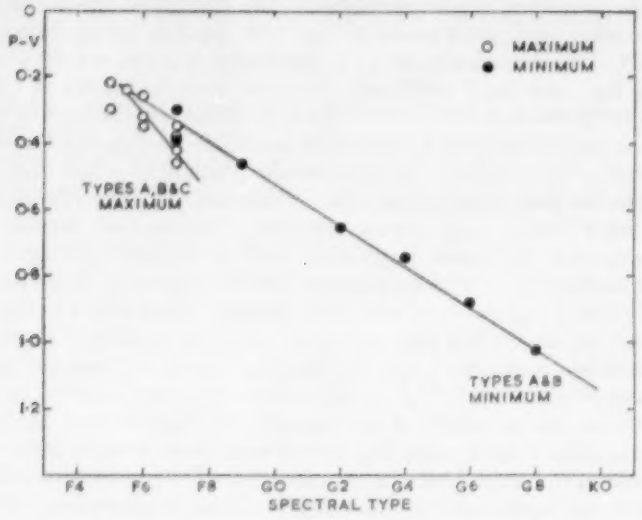


FIG. 4.—*Relations between intrinsic colour and spectral type for cepheid variables at maximum and minimum light.*

An interesting feature of the relations (12) is that the small range in spectral type which occurs at maximum light for a considerable range in period (for example, the range from $\log P = 0{\cdot}2$ to $2{\cdot}0$ only involves a change in type from F5 to F8) is accompanied by a small range in colour. In Eggen's work (**9**, p. 403) a large range in colour at maximum light was obtained, but this was due, as we have shown, to the neglect of space reddening in establishing the intrinsic colours.

4.3. *Comparison with colours of SMC cepheids.*—The **P, V** colours for Small Magellanic Cloud (SMC) cepheids at maximum light are shown in Fig. 2. Gascoigne and Kron (**11**) have drawn attention to the fact that Eggen's colours of type II cepheids in high galactic latitudes appear to be strongly related to the colours of the SMC cepheids. According to Eggen (**21**), the relation for type II cepheids is

$$(\mathbf{P}\text{–}\mathbf{V})_{\max} = +0^{m}{\cdot}08 + 0^{m}{\cdot}21 \log P. \qquad (13)$$

From equations (8) and (13) we have

$$(\mathbf{P}\text{–}\mathbf{V})_{\max}\text{: Type I } \textit{minus} \text{ Type II} = +0^{m}{\cdot}09 - 0^{m}{\cdot}03 \log P,$$

the root mean square error in each coefficient being $\pm 0^{m}{\cdot}05$ on the assumption of error-free values in equation (13). This suggests that there is also a strong relationship between the Type I and Type II colours at maximum light, particularly in the slope of the intrinsic colour relation. The present work therefore lends some support to a recent suggestion, due to Eggen (**21**), that all the cepheids may have the same colour at maximum light—both Types I and II, and that the high latitude galactic cepheids of Type II and the SMC cepheids are the only unreddened cepheids known. It may, however, be necessary to modify this suggestion if the obscuration of the extra-galactic nebulae in directions surrounding the Small Cloud in galactic coordinates $l = 268°$, $b = -45°$ is

approximately equal to the average obscuration. According to Oort (**3**, p. 237) the average photographic absorption for extra-galactic objects in galactic latitude b is given by $A_{pg}=0^{m}\cdot31\operatorname{cosec}b$. If, as seems likely, most of the obscuration is due to selective absorption by matter near the galactic plane, the average colour excess for objects in the Small Cloud is $0\cdot31\sqrt{2}/4\cdot75$ or 0·09 magnitudes. A colour absorption of this order would probably place the intrinsic colours of the SMC cepheids beyond the statistical limits set by the uncertainty in the solution adopted for the colours of Type I galactic cepheids. Coincidence between the two sets of colours would then require more space reddening than is at present indicated by the B stars. However, the interesting fact would remain that the intrinsic colour relations at maximum light for cepheids belonging to different stellar populations have approximately the same coefficient of $\log P$.

5. *Application to galactic rotation*

An examination is now made of the connection between differential galactic rotation of the nearby cepheids and the zero-point of the period-luminosity curve, using cepheid distances corrected for absorption by means of colour excesses based upon the intrinsic colours (8). This can be carried out by means of a distance scale-factor similar to that introduced by Camm (**25**) and Mineur (**26**).

5.1. *The distance scale-factor.*—Let M be the absolute magnitude of a cepheid as given by the adopted period–luminosity curve, and suppose that the true zero-point is ΔM magnitudes fainter than adopted, that is, the true absolute magnitude is $M+\Delta M$. If the distance modulus is corrected for absorption by means of a colour excess which does not depend upon ΔM, we have the following relations for the distance modulus:

$$m_0-M=5\log r-5=\Delta M+5\log r_t-5, \tag{14}$$

where r and r_t are the distances calculated for absolute magnitudes M and $M+\Delta M$ respectively. In this case we obtain

$$r_t=\lambda r, \tag{15}$$

where λ, the distance scale-factor, is related to the zero-point correction by

$$\Delta M=-5\log\lambda. \tag{16}$$

But the colour excesses based upon the intrinsic colours (8) are dependent upon the zero-point of the period-luminosity curve. Equation (14) should then be

$$m_0-M=5\log r-5=\Delta M+\chi\Delta E+5\log r_t-5, \tag{17}$$

where ΔE is the increase in colour excesses which would arise if the solution for the intrinsic colours (8) had been made with absolute magnitudes $M+\Delta M$. The zero-point correction and the distance scale-factor are then related by

$$\Delta M=\frac{-5\log\lambda}{1+\chi\Delta E/\Delta M}. \tag{18}$$

It can be shown that the coefficients of $\log P$ in the relations (8) are approximately independent of ΔM and that the intrinsic colours are increased by $0^{m}\cdot07$ per magnitude decrease in absolute brightness. If we adopt $\Delta E/\Delta M\simeq0\cdot07$, equation (18) becomes $\Delta M\simeq7\cdot5\log\lambda$.

5.2. *Oort's constant and the zero-point correction.*—The amplitude of the double-wave term in Oort's approximation to the effect of galactic rotation on the

radial velocities is proportional to the true distance r_t. Since $r_t = \lambda r$, Oort's equation may be written

$$V' = \lambda A r \sin 2(l - l_0) \cos^2 b, \tag{19}$$

where V' is the radial velocity corrected for solar motion. The solution of equation (19) by least squares gives

$$\lambda A = \frac{\Sigma V' r \sin 2(l - l_0) \cos^2 b}{\Sigma r^2 \sin^2 2(l - l_0) \cos^4 b}. \tag{20}$$

In this way we determine not A, but λA, where λ is related to the zero-point correction ΔM by equations (16) or (18).

The solution (20) assumes that the longitude of the galactic centre and the corrections for solar motion are known. We shall adopt $l_0 = 325°$, and corrections for the standard value of the local solar motion (**27**). The data on cepheid colours used in Section 3 are here augmented by provisional colour excesses for eighteen cepheids kindly made available by Eggen. Of the fifty-six stars for which colour excesses and radial velocities are available, the distances of all but four are less than 2 kpc. It was considered necessary to impose the condition $r \leqslant 2$ kpc since Oort's formula is not accurate for large distances. This reduces the sample to fifty-two stars which, although small, is quite representative of galactic cepheids in so far as the period distribution and velocity dispersion is concerned. The following solution was obtained for λA and the root mean square error:

$$\lambda A = 19{\cdot}8 \pm 2{\cdot}3, \tag{21}$$

the units being km per second per kiloparsec of true distance r_t. The root mean square error of an observational equation was found to be $\pm 10{\cdot}8$ km/sec, and the average velocity residual without regard to sign 8·8 km/sec. The mean distance of the sample $\bar{r} = 0{\cdot}87$ kpc gives an amplitude $\lambda A \bar{r} = 17{\cdot}2$ km/sec for Oort's double wave (19).

TABLE II

Oort's constant and the zero-point correction to the P–L curve

ΔM	$A = (\lambda A) \times 10^{\Delta M/\alpha}$	
	$\alpha = 5$	$\alpha = 7{\cdot}5$
mag.	km/sec/kpc	km/sec/kpc
−1·0	12·5	14·6
−0·5	15·7	17·0
0·0	19·8	19·8
+0·5	24·9	23·1
+1·0	31·4	26·9
+1·5	39·6	31·4

The connection between A and ΔM may be written

$$A = (\lambda A) \times 10^{\Delta M/\alpha}, \tag{22}$$

where $\alpha = 5$ or 7·5 according as equation (16) or (18) for ΔM is used. From equations (21) and (22) we obtain the values of A given in Table II. These

are the values which the sample would give for Oort's constant if Baade's revised zero-point of the P–L curve were corrected by ΔM magnitudes. The marked dependence upon ΔM demands that some explanation be offered for the fact that when $\Delta M = 0$ (Baade's zero-point) we recover the value $A = 20$ km/sec/kpc originally found by Joy (**28**) using Shapley's P–L curve (**29**) for which $\Delta M = +1{\cdot}5$ magnitudes relative to Baade's zero-point. If the dependence of the intrinsic colours on ΔM is neglected, the cepheid distances are halved and Oort's constant doubled for $\Delta M = +1{\cdot}5$. Thus we should obtain $A = 40$ km/sec/kpc on the basis of Shapley's zero-point, whereas Joy found $A = 20$ km/sec/kpc. The difference is due to the amount of interstellar absorption allowed for in the two investigations. In Joy's work, the model of uniform absorption was adopted, the coefficient of absorption being $0^{m}{\cdot}85$ per kpc for all galactic latitudes and longitudes. In the present work, the absorption depends upon both latitude and longitude, and it is necessary to define average values before a comparison can be made with Joy's absorption.

Following Stebbins and Huffer (**1**, p. 247) we shall consider two averages for the coefficient of selective absorption given by the colour excesses: $c_1 = \Sigma E/\Sigma r$, and $c_2 = (1/n)\Sigma E/r$. Changing to total photographic absorption, the averages for the fifty-two stars used in the above solutions are $1^{m}{\cdot}83$ and $2^{m}{\cdot}30$ respectively. We can obtain a third estimate of the average absorption by means of the B star colour excesses used in the determination of the intrinsic colours of the cepheids. Referring to the exponential model of the absorbing medium, the average absorption for objects within the distance z from the galactic plane is given by (see **4**, p. 43)

$$\overline{A_{pg}/r} = \frac{a_0}{z}\int_0^z \frac{1-e^{-z/\beta}}{z/\beta}\,dz = a_0 \sum_1^{\infty} \frac{(-1)^{n+1}}{n\,.\,n!}\left(\frac{z}{\beta}\right)^{n-1}. \qquad (23)$$

Now the average value of a_0 for the 37 stars used in the Parenago solution in Section 3 is $2^{m}{\cdot}74$ per kpc, and of the 52 stars used in the solution (20) there are 21 with $z < 50$ pc and 40 with $z < 100$ pc, the average value of z being 69 pc. From equation (23), with $\beta = 100$ for the absorbing medium, the average absorption for objects within 50 pc or 100 pc from the galactic plane is $2^{m}{\cdot}43$ per kpc or $2^{m}{\cdot}19$ per kpc respectively. These three estimates place the average absorption near 2 mag. per kpc compared with Joy's value $0^{m}{\cdot}85$ per kpc.

The recovery of Joy's original value of Oort's constant in the present work is due to the fact that the greater absorption compensates for the increase in the distance moduli introduced by Baade's correction of Shapley's zero-point. The effectiveness of this compensation within the context of uniform absorption can be seen as follows: Let $\bar{r}$ be the mean distance of the sample for Baade's zero-point and the coefficient of absorption a, $\bar{r}_1$ being the mean distance for Shapley's zero-point ΔM magnitudes fainter and for the coefficient of absorption a_1. Averaging over the stars in the sample we obtain

$$\overline{M} = \overline{m} + 5 - 5\log\bar{r} - a\bar{r} - 5C$$

$$\overline{M} + \Delta M = \overline{m} + 5 - 5\log\bar{r}_1 - a_1\bar{r}_1 - 5C_1,$$

where C and C_1 are corrections for dispersion in distance given by

$$C = \overline{\log r} - \log\bar{r}, \qquad C_1 = \overline{\log r_1} - \log\bar{r}_1.$$

The equation for the mean distance is then

$$5 \log \bar{r}_1 + a_1 \bar{r}_1 = 5(C - C_1) + 5 \log \bar{r} + a\bar{r} - \Delta M.$$

In this equation we take $\bar{r} = 0{\cdot}87$ kpc, $\Delta M = +1{\cdot}5$, $a_1 = 0^m{\cdot}85$ per kpc, and $a \simeq 2^m$ per kpc. If we neglect the term $5(C - C_1)$ on the ground that the effect of the above mentioned compensation is to equalize the dispersions, we obtain $\bar{r}_1 \simeq 0{\cdot}73$ kpc. Considering the limitations imposed by the model of uniform absorption, and the difficulty in obtaining a suitably defined average coefficient of absorption, we may conclude that the mean distance of the sample when referred to Shapley's zero-point and Joy's coefficient of absorption does not differ significantly from the mean distance referred to Baade's zero-point and the higher value of the coefficient of absorption.

Conclusion.—The intrinsic colours of Type I cepheids adopted in the paper give an average absorption of about 2 magnitudes per kpc for objects near the galactic plane, the result being dependent to some extent upon how the absorption is averaged. Evidence for the high absorption was obtained in 1938 by Oort, who found an average coefficient of selective absorption in the galactic plane of $0^m{\cdot}32$ per kpc from the colour excesses of B stars on the system of Stebbins, Huffer and Whitford. In Oort's work, a correction $+0^m{\cdot}049$ was applied to the zero-point on the SHW colour excesses and the minimum value $A_{pg}/E_1 = 5{\cdot}5$ was adopted. It now appears that the correction should not be greater than $+0^m{\cdot}04$, and the ratio $A_{pg}/E_1 = 9$ seems well-established. Oort's data then give an average coefficient of selective absorption $0^m{\cdot}26$ per kpc for objects within 100 pc of the galactic plane, and an average photographic absorption $2^m{\cdot}3$ per kpc. A similarly high value of the average absorption near the galactic plane has been advocated by Parenago and van Rhijn.

Some of the direct consequences of the high absorption are (1) the similarity between the colours of cepheid variables in the Small Magellanic Cloud and the intrinsic colours of Type I galactic cepheids, the latter being hitherto supposed to be redder, and (2) the recovery of Joy's original value of Oort's constant when Baade's correction is applied to Shapley's zero-point of the period–luminosity curve for galactic cepheids.

Acknowledgments.—The investigation has derived much benefit from discussions with Professor H. H. Plaskett to whom the statistical treatment in Section 4.1 is largely due. The writer is indebted to Dr J. B. Sykes for his generous assistance with translations from the Russian literature, and to Dr O. J. Eggen, of the Lick Observatory, who kindly made available the reduction of his photoelectric observations to the **P, V** System, and in addition communicated some unpublished data on cepheid colours.

The work described in the paper was carried out during the author's tenure of the Radcliffe Travelling Fellowship in Astronomy at the University Observatory, Oxford.

University Observatory,
Oxford.

References

(1) J. E. Stebbins and C. M. Huffer, *Washburn Obs. Publ.*, **15**, 220, 1934.
(2) J. E. Stebbins, C. M. Huffer and A. E. Whitford, *Ap. J.*, **90**, 209, 1939.
(3) J. H. Oort, *B.A.N.*, **8**, 243, 1938.
(4) N. F. Florya, *Sternberg Publ.*, **16**, 4, 1949.
(5) P. J. van Rhijn, *Groningen Publ.*, **53**, 41, 1949.
(6) A. D. Code, *Ap. J.*, **106**, 309, 1947.
(7) G. C. Badalyan, *Comm. Burakan Obs.*, No. 8, 1951.
(8) M. A. Vashakidze, *Abastumani Astrophys. Obs. Bull.*, No. 13, 87, 1953.
(9) O. J. Eggen, *Ap. J.*, **113**, 367, 1951.
(10) H. L. Johnson and W. W. Morgan, *Ap. J.*, **114**, 526, 1951.
(11) S. C. B. Gascoigne and G. E. Kron, *P.A.S.P.*, **65**, 32, 1953.
(12) J. H. Oort, *B.A.N.*, **6**, 260, 1932.
(13) E. T. R. Williams, *Ap. J.*, **75**, 386, 1932; **79**, 404, 1934.
(14) A. Hunter and E. G. Martin, *M.N.*, **100**, 669, 1940.
(15) P. P. Parenago, *A.J.* (*USSR*), **17**, No. 4, 1, 1940.
(16) P. P. Parenago, *A.J.* (*USSR*), **22**, 129, 1945.
(17) P. J. van Rhijn, *Groningen Publ.*, **53**, 9, 1949.
(18) H. Shapley, *P.N.A.S.*, **26**, 541, 1940.
(19) A. E. Whitford, *Ap. J.*, **107**, 102, 1948.
(20) F. H. Seares and M. C. Joyner, *Ap. J.*, **98**, 272, 1943.
(21) O. J. Eggen, private communication.
(22) P. P. Parenago and N. Grigorieva, *A.J.* (*USSR*), **14**, 329, 1937.
(23) H. Shapley and V. McKibben, *P.N.A.S.*, **26**, 105, 1940.
(24) H. Shapley, *P.N.A.S.*, **26**, 681, 1940.
(25) G. L. Camm, *M.N.*, **104**, 165, 1944.
(26) H. Mineur, *Ann. d'Astrophys.*, **7**, 160, 1944.
(27) F. Link, *Czech. Astr. Soc., Memoirs and Observations*, No. 9, 1948.
(28) A. H. Joy, *Ap. J.*, **89**, 356, 1939.
(29) H. Shapley, *Star Clusters*, Harvard Monograph No. 2, 135, 1930.

SPECTROPHOTOMETRIC MEASUREMENTS OF EARLY-TYPE STARS

II. Results for stars of type B1*

W. M. H. Greaves, E. A. Baker and R. Wilson

(Communicated by the Astronomer Royal for Scotland)

Summary

The Edinburgh spectrophotometric programme aims at the measurement of equivalent widths of stellar spectral lines with the highest accuracy reasonably attainable using the available equipment. In this paper a general description is given of the technique which has been developed; it supplements the more detailed description previously published. All the lines measured lie in the spectral range 4100 A to 6800 A over which tolerable focus can be secured with the spectrograph used. For each star the equivalent widths given depend on the mean of five or six negatives and each negative has been measured at about 1500 fixed wave-lengths. Experience has confirmed that the methods used achieve a substantial gain in accuracy. Systematic errors in calibration have been reduced to the order of one per cent. Plate grain fluctuations have been reduced by combining the measures for the several negatives of each star. A further reduction in accidental fluctuations can be secured by averaging for several stars of the same type, and the formation of a mean plot for five B1 stars (27 negatives), when combined with the information given in Moore's Tables, has made it easier to choose continuous spectrum wave-lengths clear of weak lines with central depressions of the order of one per cent. In this way the falsification of the interpolated continuum which can result from the presence of weak lines is reduced. On the mean B1 plot the standard errors of measurement at one wave-length are $\pm 0{\cdot}008$ and $\pm 0{\cdot}006$ magnitudes in the blue and red respectively. This mean B1 plot also reveals the presence of faint blending lines and a reasonably satisfactory treatment of blends becomes possible.

Equivalent widths of 28 spectral lines are given for each of seven stars. For each line the limits of integration used and the standard error of the equivalent width are also given. The standard error depends, *inter alia*, on the range of integration; for the weaker lines the s.e. is about $\pm 0{\cdot}07$ equivalent angstroms.

For lines other than hydrogen or helium there is a systematic dependence of equivalent width on stellar luminosity, the equivalent widths being larger for the higher luminosities. This progression covers a range of more than two to one in equivalent width. For the Balmer lines Hα, Hβ and Hγ the progression is reversed, the equivalent widths being smaller for the higher luminosities.

The results for the He I lines are of special interest. When the singlet and triplet lines 6678 and 5876 are included in a discussion of the material, it becomes impossible to reconcile the equivalent widths with the theoretical curve of growth. There is a deficiency in the measured equivalent widths of 6678 and 5876 which seems to increase with decreasing luminosity. It is suggested that cyclical transition processes are operative, and are more pronounced for the lower luminosities.

The Balmer lines of β Cephei have been found to be variable, the variation being especially pronounced for Hα. From the present material no statement can be made as to whether the variation is periodic or irregular.

* The full text of this paper appears in *Publications of the Royal Observatory, Edinburgh*, 1, No. 6, 1955.

CONTENTS

PAGE

Meeting of 1955 March 11:
- Fellow elected ... 219
- Presents announced ... 219

Meeting of 1955 April 6:
- Fellows elected ... 219
- Associate elected ... 219
- Presents announced ... 219
- Presentation of the Gold Medal ... 220

Dirk Brouwer, The motions of the outer planets (George Darwin Lecture) ... 221

A. D. Thackeray, RR Telescopii in 1953 and 1954, and the development of slow novae ... 236

E. M. Lindsay, Emission objects in the Small Magellanic Cloud showing the N_1, N_2 nebular lines... ... 248

H. H. Plaskett, Interpretation of Fraunhofer-line profiles ... 256

W. R. Hindmarsh, An interferometric investigation of the effect of hyperfine structure on solar line profiles ... 270

M. J. Seaton, Continuum intensities in planetary nebulae ... 279

A. E. Roy and M. W. Ovenden, On the occurrence of commensurable mean motions in the solar system. The mirror theorem ... 296

W. B. Bonnor, The instability of the Einstein universe ... 310

D. W. N. Stibbs, The intrinsic colours of galactic cepheid variables ... 323

Summary of paper published in the Publications of the Royal Observatory, Edinburgh

W. M. H. Greaves, E. A. Baker and R. Wilson, Spectrophotometric measurements of early type stars. II. Results for stars of type B1 ... 338

Printed in London by Taylor & Francis, Ltd., 1956 April